

**Entwicklung und Evaluation eines Verfahrens zur
Lernverlaufsdokumentation im Lernbereich Rechtschreiben
der 2. Klassenstufe**

**Dissertation
zur
Erlangung des akademischen Grades
Doctor philosophiae (Dr. phil.)
der Philosophischen Fakultät
der Universität Rostock**

vorgelegt von

**Kristin Kuhlmann, geb. am 31.10.1983 in Neubrandenburg
aus Köln**

Rostock, 02.01.2013

Verteidigung am 13.12.2013

1. Gutachter: Prof. Dr. Bodo Hartke
Universität Rostock
Philosophische Fakultät
Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation
2. Gutachter: Prof. Dr. Katja Koch
Universität Rostock
Philosophische Fakultät
Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation
3. Gutachter: Prof. Dr. Vera Moser
Humboldt-Universität zu Berlin
Philosophische Fakultät IV
Pädagogik bei Beeinträchtigungen des Lernens und Allgemeine
Rehabilitationspädagogik

Danksagung

Mit diesen Zeilen möchte ich allen danken, die mich in den letzten Jahren meiner Dissertation begleitet und unterstützt haben.

Ich bedanke mich bei Prof. Dr. Bodo Hartke für die Themenstellung, die Betreuung der Arbeit und die Zeit, die er sich dafür nahm. Ein besonderer Dank gilt dem Cusanuswerk, das mir durch die finanzielle Förderung unabhängiges Arbeiten und Forschen ermöglichte. Ein herzliches Dankeschön auch für die ideelle Förderung, die eine Bereicherung in vielerlei Hinsicht war.

Dem Schulamt Neubrandenburg danke ich für die Zustimmung zu meiner Untersuchung. Bei den an der Untersuchung teilgenommenen Schulen, Lehrkräften und Schülern bedanke ich mich für Ihr Vertrauen und Ihre Zeit.

Für die große Unterstützung bei der Erhebung der Daten möchte ich den Studenten Ina Czub, Marina Herrmann, Marja Kubbutat, Stefanie Kuhn, Sandra und Katja Stolzenburg der Universität Rostock sowie den Studenten Anne Martin, Elisa Pasewaldt, Susanne Polgar und Lars Timer der Hochschule Neubrandenburg ganz besonders danken. Ohne ihre Unterstützung hätte ich die Untersuchung nicht realisieren können.

Dr. Yvonne Blumenthal möchte ich für die vielen und geduldigen Hinweise zu meiner Arbeit und zur Datenauswertung sowie für die aufmunternden Telefonate Danke sagen. Für die weitere hilfreiche Unterstützung bedanke ich mich bei Prof. Dr. Kirsten Diehl, Dr. Kathrin Mahlau und Stefan Voß. Für die gründlichen orthographischen und strukturellen Hinweise zu meiner Arbeit danke ich meiner Freundin Franziska Sperfeld sowie Janina Ferrari und Prof. Dr. Sabine Kuhlmann.

Meinen Freundinnen Anne-Mareike Fahrenwaldt und Stefanie Otto bin ich zu besonderem Dank verpflichtet. Sie begleiteten die vielen und manchmal unendlich erscheinenden Schritte meiner Dissertation immer so, dass ich auch in kraftraubenden Zeiten wieder lachen konnte.

An letzter Stelle, aber eigentlich zuallererst, bin ich dankbar für meine Familie, insbesondere meine Eltern und meinen Mann Arne. Sie unterstützten mich mit ihrem Zuspruch, ihrer Aufmunterung und großen Geduld immer in meinem Vorhaben, sind nun aber auch froh, dass die Dissertation abgeschlossen ist. Dem jüngsten und wichtigsten Menschen, meinem Sohn Theo, bin ich dankbar dafür, dass er mir mit seinem Lächeln immer wieder zeigt, was wirklich wichtig im Leben ist.

Inhalt

1	Einleitung.....	1
1.1	Ziel der Arbeit	1
1.2	Gegenstandsbestimmung und Fragestellung.....	6
1.3	Vorgehen und Aufbau der Arbeit.....	11
2	Diagnostik und Prävention von Lernschwierigkeiten im Aptitude- Treatment Interaction- und Response to Intervention-Ansatz	13
2.1	Zur Bedeutung der evaluativen Diagnostik.....	13
2.2	Zur Bedeutung der schulischen Prävention.....	15
2.3	Aptitude-Treatment Interaction.....	17
2.3.1	Begriffsbestimmungen.....	18
2.3.2	Entstehung des Ansatzes.....	20
2.3.3	Ziele und Funktionen des Ansatzes	22
2.3.4	Modelle der Adaption	23
2.3.5	Kritik des ATI-Ansatzes	26
2.4	Response to Intervention (RTI).....	28
2.4.1	Begriffsbestimmungen.....	28
2.4.2	Entstehung des Ansatzes.....	30
2.4.3	Ziele des Ansatzes	32
2.4.4	Aufbau und Interventionsebenen des RTI-Modells.....	33
2.4.5	Verfahren zur Lernverlaufsdokumentation: Mastery Measurement und General Outcome Measurement	36
2.4.6	Kritik des RTI-Ansatzes	45
2.5	Zusammenfassung und Schlussfolgerungen	47
3	Theoretische Grundlagen des Rechtschreiberwerbs und curriculare Vorgaben für den Lernbereich Rechtschreiben	50
3.1	Begriffsbestimmungen	50
3.2	Linguistische Grundlagen des Schriftspracherwerbs	52
3.2.1	Aufbau des deutschen Sprach- und Schriftsystems	53
3.2.1.1	Sprachliche Einheiten der unilateralen Ebenen.....	54
3.2.1.2	Sprachliche Einheiten der bilateralen Ebenen.....	57
3.2.2	Prinzipien und Regeln der konventionellen Orthographie	60
3.2.3	Zusammenfassung und Schlussfolgerungen.....	62
3.3	Modelle der Rechtschreibentwicklung.....	63
3.3.1	Das Sechsstufenmodell von Frith (1985, 1986) und das Fünf-Phasen- Modell von K. B. Günther (1986)	64
3.3.2	Das Entwicklungsmodell des Rechtschreibkönnens von May (2002)	69
3.3.3	Das Stufenmodell der Rechtschreibentwicklung von Scheerer-Neumann (2008).....	74
3.3.4	Das Kompetenzentwicklungsmodell des Rechtschreibens von Löffler und Meyer-Schepers (2008)	78

3.3.5 Zusammenfassung und Schlussfolgerungen.....	81
3.4 Modelle des Rechtschreibprozesses.....	84
3.4.1 Das Generator-Test-Prozessmodell von D. P. Simon und H. A. Simon (1973)	85
3.4.2 Zwei-Wege-Modelle des Rechtschreibens	87
3.4.2.1 Das Zwei-Wege-Modell von Barry (1994)	87
3.4.2.2 Das Zwei-Komponenten-Modell von Scheerer-Neumann (2003).....	92
3.4.3 Netzwerkmodelle.....	94
3.4.3.1 Das Netzwerkmodell des Rechtschreibens von Brown und Loosemore (1994)	95
3.4.3.2 Das Rechtschreibprozessmodell NETspell von Olson und Caramazza (1994)	97
3.4.4 Zusammenfassung und Schlussfolgerungen.....	99
3.5 Anforderungen und Inhalte des Lernbereichs Rechtschreiben in den Bildungsstandards der KMK (2005), dem Rahmenplan von Mecklenburg- Vorpommern (2009a) und den Sprachbüchern.....	102
3.5.1 Rechtschreiben in den Bildungsstandards der KMK (2005)	102
3.5.2 Rechtschreiben im Rahmenplan von Mecklenburg-Vorpommern (2004)	103
3.5.3 Rechtschreiben in Sprachbüchern	105
3.5.4 Zusammenfassung und Schlussfolgerungen.....	111
4 Methodik der Untersuchung.....	113
4.1 Untersuchungsplan.....	113
4.2 Untersuchungsinstrumente.....	116
4.2.1 Deutscher Rechtschreibtest für das erste und zweite Schuljahr von Stock und Schneider (2008a).....	117
4.2.2 Entwicklung und Erprobung des Verfahrens „Formative Erfassung der Rechtschreibleistung im 2. Schuljahr (FE-RS 2)“	119
4.2.3 Entwicklung des Lehrerfragebogens zur Einschätzung der diagnostischen Verfahren zur Lernverlaufsdokumentation und Entwicklung des Lehrerfragebogens zur Einschätzung der Rechtschreibleistung der Schüler.....	127
4.3 Stichprobengewinnung	128
4.4 Untersuchungsdurchführung und Datenauswertung.....	129
4.5 Forschungsfragen und Hypothesen.....	131
4.6 Statistische Methoden der Datenanalyse	136
5 Darstellung der Ergebnisse.....	148
5.1 Stichprobenbeschreibung.....	148
5.2 Darstellung der Ergebnisse der FE-RS 2	149
5.2.1 Differenzierungsfähigkeit.....	149
5.2.1.1 Häufigkeitsverteilungen, Mittelwerte und Standardabweichungen der Form A	150
5.2.1.2 Häufigkeitsverteilungen, Mittelwerte und Standardabweichungen der Form B	158
5.2.1.3 Itemschwierigkeiten und Trennschärfen	167

5.2.2 Reliabilität.....	170
5.2.2.1 Retestreliabilität.....	170
5.2.2.2 Testhalbierungsreliabilität.....	172
5.2.2.3 Interne Konsistenz.....	173
5.2.3 Validität.....	174
5.2.3.1 Inhaltsvalidität.....	174
5.2.3.2 Konstruktvalidität.....	174
5.2.3.2.1 Konkurrente Validität.....	174
5.2.3.2.2 Diskriminante Validität.....	176
5.2.3.2.3 Faktorielle Validität der Form A.....	177
5.2.3.2.4 Faktorielle Validität der Form B.....	181
5.2.3.3 Übereinstimmungsvalidität.....	185
5.2.3.4 Ergebnisse der Regressionsanalyse zur Schätzung der prognostischen Validität.....	186
5.2.3.5 Ergebnisse der klassifikatorischen Analyse zur Schätzung der prognostischen Validität.....	189
5.2.3.5.1 Prognostische Validität der Form A.....	191
5.2.3.5.2 Prognostische Validität der Form B.....	196
5.2.4 Äquivalenz der Parallelversionen.....	203
5.2.5 Veränderungssensitivität.....	211
5.2.5.1 Häufigkeitsveränderungen und Mittelwertdifferenzen der Form A...	211
5.2.5.2 Häufigkeitsveränderungen und Mittelwertdifferenzen der Form B...	215
5.3 Darstellung der Ergebnisse der Lehrerbefragung.....	220
6 Diskussion.....	223
6.1 Zusammenfassung und Diskussion der Ergebnisse.....	223
6.2 Fazit und Ausblick.....	232
Abstract.....	239
Literaturverzeichnis.....	241
Abkürzungsverzeichnis.....	263
Tabellenverzeichnis.....	267
Abbildungsverzeichnis.....	273
Eidesstattliche Versicherung.....	275
Anhang.....	277

1 Einleitung

1.1 Ziel der Arbeit

„Wer nämlich mit h schreibt, ist dämlich.“ Sprüche wie dieser sind nicht nur Merksätze, sondern enthalten Wertungen und Einstellungen. Personen, die nicht orthographisch richtig schreiben können, fallen negativ auf (Gallmann & Sitta, 1998), weil das Rechtschreibkönnen in unserer Gesellschaft einen hohen Stellenwert besitzt. Sowohl im privaten Bereich als auch in der Schule und im Beruf ist korrektes Rechtschreiben hochgeschätzt. Zugleich ist es wichtig für den schulischen und beruflichen Erfolg, denn Rechtschreibleistungen werden für die Bewertung anderer Schulleistungen mit einbezogen. Kemmler wies 1976 nach, dass Rechtschreibleistungen einen großen Einfluss auf die Entscheidung haben, welche weiterführende Schule besucht wird. Auch knapp 20 Jahre später ließ sich ein Zusammenhang zwischen den schriftlichen Schulleistungen am Ende der 4. Klassenstufe und der Wahl der Schulart im Sekundarbereich belegen (Klicpera, Gasteiger-Klicpera & Schabmann, 1993). Heute relativieren sich diese Aussagen, da bekannt ist, dass weitere Merkmale und individuelle Leistungen, wie z. B. Intelligenz, Leseverständnis oder mathematische Leistungen, Einfluss auf spätere Schullaufbahnentscheidungen haben (P. Marx, 2007). Dennoch, Rechtschreiben ist nach wie vor eine Schlüsselkompetenz und ein wichtiges Element schriftsprachlichen Handelns (Augst & Dehn, 2007; Scheerer-Neumann, 2008).

Vielen Schülern bereitet aber das Rechtschreiben trotz häuslicher und schulischer Übungen Mühe. Vermutlich ist Rechtschreiben auch deshalb das unbeliebteste Schulfach in der Grundschule und gilt wie vor 100 Jahren als Schulkreuz (Valtin, Meyer-Schepers & Löffler, 2003; Valtin & Wagner, 2002).

„Die deutsche Rechtschreibung nämlich ist nichts weiter als ein wahres Schulkreuz; denn wenn man die Zeit, die dafür aufgewendet wird, den Ärger, den sie Eltern und Lehrern bereitet, die Tränen, die um ihretwillen von den Schülern alljährlich vergossen werden, summieren könnte, man würde erschrecken über das Unheil, das dieser Unterrichtsgegenstand Jahr für Jahr anrichtet“ (Kosog, 1912, S. 3).

Auch wenn Rechtschreibfehler sowohl bei guten als auch bei schwachen Schülern vorkommen, treten bei einigen Schülern große Schwierigkeiten beim Rechtschreiberwerb auf. Sie begehen deutlich mehr Rechtschreibfehler, Phonem-Graphem-Fehler, Wortauslassungen und willkürliche Schreibungen und haben Schwierigkeiten in der

Anwendung orthographischer Regelmäßigkeiten (Klicpera et al., 1993). Wie viele Schüler als rechtschreibschwach gelten, ist nicht eindeutig bekannt (Scheerer-Neumann, 2008). Untersuchungen zeigen aber, dass Jungen häufiger als Mädchen betroffen sind (Gasteiger-Klicpera & Klicpera, 2004).

Häufig manifestieren sich diese Rechtschreibschwierigkeiten und können bis in das Erwachsenenalter persistieren, wenn ihnen nicht entgegengewirkt wird. Längsschnittliche Analysen (Grube & Hasselhorn, 2006; Klicpera et al., 1993; P. Marx, 2007; Schneider, Stefanek & Dotzler, 1997) zeigen, dass schriftsprachliche Leistungen ab der 1. und 2. Klassenstufe relativ stabil bleiben. Schüler, die nach den ersten zehn Schulwochen sehr schwache Leistungen im Lesen und Rechtschreiben haben, gehören mit hoher Wahrscheinlichkeit auch am Ende der 1. Klassenstufe zu den Leistungsschwächeren. Die Stabilität der Rangplätze zeigt sich auch in der weiteren Entwicklung. Schüler, die am Ende der 1. bzw. in der Mitte der 2. Klassenstufe Schwierigkeiten im Lesen und Rechtschreiben haben, gehören trotz Lernzuwächsen auch am Ende der 4. und 8. Klassenstufe überwiegend zu den schwachen Lesern und Rechtschreibern (Klicpera et al., 1993). Auch im Alter von 17 bis 23 Jahren setzt sich die Stabilität der Rechtschreibleistungen fort. Die Rechtschreibleistungen am Ende der Grundschulzeit sagen zu 50 % die Leistungen im Alter von 23 Jahren voraus (Schneider & Stefanek, 2007). Dennoch sind individuelle Veränderungen innerhalb der Rangplätze möglich (Schneider et al., 1997). Insgesamt sind aber die Unterschiede zwischen Leistungsgruppen mittel- und langfristig recht stabil, sodass sich Rechtschreibschwierigkeiten nicht nur auf die Schullaufbahn, sondern auch auf die berufliche Entwicklung auswirken.

Darüber hinaus haben Rechtschreibschwierigkeiten Auswirkungen auf die emotionale und psychosoziale Entwicklung. Lese- und rechtschreibschwache Schüler, vor allem Jungen, entwickeln häufig ein niedriges Selbstkonzept. Folgen dieser geringen Selbsteinschätzung können zunehmende Lernrückstände, aggressives oder depressives Verhalten, Schulangst, geringe Motivation oder Vermeidungsverhalten sein. Durch das überwiegend internalisierende Verhalten werden sie von Eltern, Lehrern und Mitschülern als auffällig beurteilt. Sie gelten als zurückgezogen, ängstlich und scheu. Da sie aufgrund ihres Verhaltens von Mitschülern geschnitten werden, geraten Schüler mit Lese- und Rechtschreibschwierigkeiten in Außenseiterpositionen. Sie leiden unter ihrer Situation, sind überwiegend einsam und ziehen sich zurück (Gasteiger-Klicpera, Klicpera &

Schabmann, 2006; Schulz, Dertmann & Jagla, 2003).

In Anbetracht der Auswirkungen von Rechtschreibleistungen auf die schulische und berufliche Zukunft und des Leidensweges von Schülern mit Rechtschreibschwierigkeiten ist es eine dringende Aufgabe, Schüler mit Schwierigkeiten im Rechtschreiberwerb so früh wie möglich zu identifizieren und entsprechend zu fördern. Fehlentwicklungen sollten bereits in ihrer Entstehung, das heißt im Vorschulalter und im Anfangsunterricht, erkannt werden, um eine frühe Förderung gewährleisten zu können (Barth & Gomm, 2008). Zur Erfassung der Rechtschreibleistung von Schülern steht eine Vielzahl von normierten Rechtschreibtests zur Verfügung. In der deutschen Testzentrale werden sieben Verfahren zur Erfassung der Rechtschreibleistungen in 2. Klassen aufgeführt:

1. der Deutsche Rechtschreibtest für das erste und zweite Schuljahr (DERET 1-2+) (Stock & Schneider, 2008a),
2. der Diagnostische Rechtschreibtest für zweite Klassen (DRT 2) (R. Müller, 2004),
3. die Hamburger Schreib-Probe für die Klassenstufen 1/2 (HSP 1+) (May, 2005a) und für die Klasse 2 (HSP 2) (May, 2005b),
4. die Münsteraner Rechtschreibanalyse (MRA) (Schönweiss, 2008) und
5. der Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen (WRT 1+) (Birkel, 2007a) sowie für zweite und dritte Klassen (WRT 2+) (Birkel, 2007b).

Daneben werden in der Testzentrale Mehrfächertests mit einem Rechtschreibteil und testunabhängige fehleranalytische Verfahren für die 2. Klassenstufe aufgeführt. Zu den Mehrfächertests mit einem Rechtschreibteil gehören:

1. der Allgemeine Schulleistungstest für zweite Klassen (AST 2) (Rieder & Ingenkamp, 1991),
2. der Salzburger Lese- und Rechtschreibtest (SLRT) (Landerl, Wimmer & Moser, 2006) und
3. die Schultestbatterie zur Erfassung des Lernstandes in Mathematik, Lesen und Schreiben II (SBL II) (Kautter, Storz & Munz, 2002).

Zu den fehleranalytischen Verfahren, die ergänzend zu Rechtschreibtests und freien Texten herangezogen werden können und eine differenzierte Rechtschreibfehleranalyse erlauben, gehören:

1. die Dortmunder Rechtschreibfehler Analyse (DoRA) (Löffler & Meyer-Schepers, 1992),

2. die Aachener Förderdiagnostische Rechtschreifehler-Analyse (AFRA) (Herné & Naumann, 2002) sowie
3. die Oldenburger Fehleranalyse (OLFA) (G. Thomé & D. Thomé, 2004b), die aber erst ab der 3. Klassenstufe anwendbar ist.

Aufgrund der großen Auswahl von Rechtschreibtests und fehleranalytischen Verfahren scheint die diagnostische Erfassung der Rechtschreibleistungen leicht. „Die verlässliche Erfassung der Rechtschreibfähigkeit ist [aber] aus verschiedenen Gründen nur begrenzt möglich“ (Hemminger, Roth, Schneck, Jans & Warnke, 2000, S. 189). Die erste Problematik betrifft das Rechtschreiben an sich, denn Rechtschreibleistungen sind von individuellen Tagesschwankungen und dem Unterricht abhängig. Zudem ist das Maß der Rechtschreibleistung von der Art und Weise der diagnostischen Überprüfung, das heißt von der Art des Anforderungsprofils, abhängig. Ein weiteres Problem ist, dass es keinen einheitlichen, bundesweiten Grundwortschatz gibt (Hemminger et al., 2000). Die zweite Problematik betrifft die Konzeption normbasierter Rechtschreibtests. Lehrkräfte können mit Hilfe normbasierter Rechtschreibtests die Rechtschreibleistungen eines Schülers oder einer Klasse mit einer Normgruppe vergleichen und durch die Abweichung der Rechtschreibleistungen von dieser Norm die im Test erzielten Leistungen schnell und leicht bewerten. Normbasierte Leistungstests können aber nicht alle curricularen und regionalen Besonderheiten, wie z. B. Lehrpläne, Schulbücher oder Stoffverteilungspläne, berücksichtigen, wodurch die im Test erzielten Leistungen falsch bewertet werden könnten. Ein weiteres Problem ist, dass die Einschätzung der Rechtschreibleistung aus Gründen der Ökonomie nur auf Basis weniger Wörter und Testitems vorgenommen wird. Das heißt, zu den jeweiligen Messzeitpunkten decken Rechtschreibtests die wichtigsten, aber nicht alle Bereiche der Orthographie ab. Des Weiteren erfassen Rechtschreibtests die Leistungen oft nur zu Beginn oder am Ende einer Klassenstufe, wodurch eine zeitlich engmaschige und kontinuierliche Dokumentation der Lernentwicklung und des Lernverlaufs nicht möglich ist. Dadurch bleibt für Lehrkräfte unklar, welche Lernschritte Schüler bereits bewältigt haben und welche nicht. Zudem besteht die Gefahr, Rechtschreibschwierigkeiten nicht rechtzeitig erkennen und entsprechende Interventionen durchführen zu können. Darüber hinaus erfasst bislang nur die HSP 1+ (May, 2005a) frühzeitig, das heißt Mitte der 1. Klassenstufe, die Rechtschreibleistungen. Andere Verfahren wie der WRT 1+ (Birkel, 2007a) sind erst ab dem Ende der 1. Klassenstufe anwendbar, wodurch der Anspruch einer möglichst frühzeitigen Diagnostik und Förderung

eingeschränkt wird. Der letzte Problembereich betrifft die mangelnde theoretische Fundierung einiger Rechtschreibtests, z. B. der DRT 2 (R. Müller, 2004), der WRT 1+ (Birkel, 2007a), der WRT 2+ (Birkel, 2007b) und der AST 2 (Riedel, 1991). Diese gründen nicht auf Modellen des Rechtschreibprozesses und der Rechtschreibentwicklung, wobei gerade entwicklungs- und prozesstheoretische Modelle mögliche Erklärungsansätze für Schreibungen und Informationen für pädagogisches Handeln bieten (Jansen, 2008; Walter, 2009b).

Die bisherigen Ausführungen zeigen, dass mit der Diagnostik von Rechtschreibleistungen mehrere Schwierigkeiten verbunden sind. Zugleich ist deutlich geworden, dass im Bereich der Rechtschreibdiagnostik eine Forschungslücke besteht, da es bislang kein diagnostisches Verfahren gibt, dass die Rechtschreibleistungen frühzeitig, zeitlich engmaschig, veränderungssensitiv und theoretisch fundiert erfasst. Diese Forschungslücke soll die vorliegende Arbeit verringern. Die Arbeit soll einen Beitrag zur frühen, entwicklungsbezogenen und lernprozessbegleitenden Diagnostik von Rechtschreibleistungen leisten.

Ziel dieser Arbeit ist die Entwicklung und Evaluation eines curriculumvaliden Verfahrens zur Erfassung der Rechtschreibentwicklung, der „Formativen Erfassung der Rechtschreibleistung im 2. Schuljahr (FE-RS 2)“. Ziel der FE-RS 2 ist es, durch wiederholte, kontinuierliche Messungen die Lernentwicklung und den Lernfortschritt im Rechtschreiberwerb zu dokumentieren. Auf Grundlage theoretischer Modelle zum Rechtschreiberwerb und Rechtschreibprozess soll damit beurteilt werden können, welche Lernschritte bewältigt wurden und welche nicht. Veränderungen im Sinne von individuellen Lernzuwachsen oder Lernstagnationen im Rechtschreiberwerb sollen sichtbar werden. Das Verfahren zur Erfassung der Rechtschreibentwicklung wird bewusst für die 2. und nicht für die 1. Klassenstufe der Grundschule konzipiert. Grund dafür ist die noch mangelnde Fähigkeit der Schüler 1. Klassen, das Geschriebene rechtschriftlich zu kontrollieren.

„Sie [die Kontrolle des Geschriebenen] bildet sich im Laufe des Lesen- und Schreibenlernens erst in der 2. Klasse aus. Dann erst sind die ersten Kinder sichere Leser und können aufgrund dieser Qualität der Lesefertigkeit sicher herausfinden, dass sie, wenn sie etwas ‚Falsches‘ auf der Basis des selbst Geschriebenen gelesen haben, es auch tatsächlich falsch geschrieben haben. Ein junger Mensch in der 1. Klasse kann nicht wissen, wenn er etwas falsch gelesen hat, ob das daran liegt, dass er falsch

geschrieben und richtig gelesen hat oder ob er das richtig Geschriebene falsch gelesen hat“ (Mannhaupt, 2008, S. 138).

In Anlehnung an Mannhaupt (2008) wird davon ausgegangen, dass ein gezielter Rechtschreibunterricht erst in der 2. Klasse nach Abschluss des Erstleselehrganges sinnvoll ist. Innerhalb des Erstleseunterrichts ist zuvor das lautorientierte Schreiben zu üben. Erste Einsichten in die orthographischen Regelmäßigkeiten und Strukturen gewinnen Schüler erst nach der alphabetischen Phase, also in der 2. Klasse nach Abschluss des Erstleselehrganges (Scheerer-Neumann, 2007, 2008).

Das frühzeitige Erkennen von Entwicklungsverläufen rechtschriftlicher Fähigkeiten soll „einen ersten Beitrag zur Verhinderung von Minderleistungen“ (Diehl & Hartke, 2007, S. 209) leisten. Zugleich soll die FE-RS 2 Hinweise für die Planung der Förderung der Rechtschreibleistungen bieten und deren Evaluation ermöglichen.

1.2 Gegenstandsbestimmung und Fragestellung

Die *Lernverlaufsdiagnostik* ist ein Ansatz, „die Lernentwicklung und den Lernfortschritt einzelner Kinder wie ganzer Klassen über längere Zeit hinweg zu dokumentieren“ (Klauer, 2011, S. 208). Bei der Lernverlaufsdiagnostik stehen anders als bei gängigen punktuellen Schulleistungstests oder lehrzielorientierten Tests nicht die Selektion, Klassifikation oder Platzierung der Schüler oder die Lehrzielüberprüfung im Mittelpunkt. Im Mittelpunkt der Lernverlaufsdiagnostik stehen die Erfassung und Dokumentation der individuellen *Leistungsveränderungen in einer gleichbleibenden Situation im Verlauf der Zeit* (Klauer, 2006, 2011; Wilbert & Linnemann, 2011). Verfahren der Lernverlaufsdiagnostik sind somit *formative* Verfahren, die parallel zum Unterricht eingesetzt werden und die Lernentwicklung und den Lernverlauf der Schüler erfassen. Sie beantworten Fragen wie „Was hat der Schüler unter dem Einfluss des gegenwärtigen Unterrichts und von aktuell praktizierten Fördermaßnahmen bereits gelernt? Wo hat er noch Schwierigkeiten?“ (Diehl & Hartke, 2011, S. 122). Um die Lernentwicklung und Leistungsveränderungen in einem Lernbereich erfassen zu können, bearbeiten Schüler in kurzen, regelmäßigen Zeitabständen Aufgaben zu einem Lehrziel. Auf diese Weise lässt sich dokumentieren, wie sich die Leistungen in einem Lernbereich über einen längeren Zeitraum entwickeln (Klauer, 2006, 2011; Wilbert & Linnemann, 2011). Zudem geben die Lernkurven den Lehrkräften

Rückmeldungen über die Effektivität des Unterrichts. Die Lernkurven geben auch Hinweise darüber, welche Schüler besondere Fördermaßnahmen benötigen. Insgesamt hat die Lernverlaufsdiagnostik viel Potenzial, wenn Lehrkräfte die diagnostischen Informationen gezielt zur Anpassung des Unterrichts und zur Ableitung von Fördermaßnahmen nutzen (Souvignier & Förster, 2011). Sie kann sodann „Lehrkräfte dabei unterstützen, angemessene Förderkonzepte zu realisieren und so höhere Lernzuwächse zu bewirken“ (Souvignier & Förster, 2011, S. 244).

In den Niederlanden und den USA wurden effektive Verfahren der Lernverlaufsdiagnostik entwickelt, um die Lernentwicklung zu dokumentieren und um Unterricht und Förderung datenbasiert zu optimieren. In den Niederlanden wurde das *Leerlingonderwijsvolgsysteem* (LOVS) und in den USA wurden *curriculumbasierte Messverfahren* entwickelt. Das LOVS ist ein theoriebasiertes Verfahren zur *langfristigen* Dokumentation des Lernverlaufs in unterschiedlichen Lernbereichen und wird pro Lernbereich ca. zweimal im Schuljahr eingesetzt. Curriculumbasierte Verfahren nutzen die im Unterricht verwendeten Materialien und werden *in kurzen Zeitabständen* wiederholt im Schuljahr eingesetzt. Im Vergleich zum LOVS sind curriculumbasierte Verfahren damit sensibler für Veränderungen (Diehl & Hartke, 2007), weshalb sich die vorliegende Arbeit in der Entwicklung der FE-RS 2 vorwiegend auf US-amerikanische Verfahren der Lernverlaufsdiagnostik stützt.

In den USA liegen die frühesten Forschungsanstrengungen zu datenbasierten Unterrichts Anpassungen mit Hilfe von Messverfahren, die die Reaktionen der Schüler auf den Unterricht und Interventionen erheben (Response to Intervention-Ansatz) sowie den Lernverlauf dokumentieren, in den 1960er Jahren. Aber erst in den letzten Jahrzehnten nahmen Diskussionen zur Notwendigkeit datenbasierten Unterrichts und veränderter Diagnostik von Lernschwierigkeiten in Forschungseinrichtungen und Schulen zu. In den USA werden diese Forschungsanstrengungen durch die Gründung neuer Kommissionen und den Erlass neuer Gesetze unterstützt, die für diesen neuen datenbasierten Prozess und die Verbesserung der sonderpädagogischen Förderung stimmen (Bradley, Danielson & Doolittle, 2007; Shores & Bender, 2007; Senate and House of Representatives of the United States of America in Congress, 2004; U.S. Department of Education, 2002). In Deutschland begann die Auseinandersetzung mit Verfahren der Lernverlaufsdiagnostik durch die Überblicksartikel von Klauer (2006) sowie Diehl und Hartke (2007). Walter

(2008b, 2009a) überprüfte daraufhin unterschiedliche Techniken der Lernverlaufsdagnostik für die Lesekompetenz und konzipierte für den Schul- und Förderbereich das curriculumbasierte Verfahren „Lernfortschrittsdiagnostik Lesen“ (2010). Auch Diehl und Hartke (2012) entwickelten ein curriculumbasiertes Verfahren zur Abbildung des Lernverlaufs im Lesen, um durch wiederholte Messungen die Leistungsentwicklung des Lesens zu erfassen und bei Lernstagnationen oder -rückläufen Förderung anzuregen. Strathmann und Klauer (2008) sowie Strathmann, Klauer und Greisbach (2010) entwickeln innerhalb einer Pilotstudie zurzeit ein Verfahren der Lernverlaufsdagnostik für den Lernbereich Rechtschreiben in der Grundschule. Darüber hinaus wird von Strathmann und Klauer (in Vorbereitung) ein Verfahren der Lernverlaufsdagnostik für den Mathematikunterricht in 2., 3. und 4. Grundschulklassen vorbereitet. Knopp und Hartke (2010) konzipierten mit dem Inventar „Rechenfische“ ebenfalls ein Verfahren der Lernverlaufsdagnostik für den Mathematikunterricht in der 1. Klassenstufe. Internetbasierte Verfahren der Lernverlaufsdagnostik zur Lesekompetenz der 3. bis 6. Klassenstufe, zur mathematischen Kompetenz der 1. bis 3. Klassenstufe und für den Englischunterricht der 5. Klasse (www.quop.de) wurden in einem Forschungsprojekt von Souvignier entwickelt (Westfälische Wilhelms-Universität Münster, n.d.).

Mit der Konstruktion und Evaluation von Verfahren der Lernverlaufsdagnostik sind besondere Anforderungen verbunden. Die Anforderungen entsprechen im Wesentlichen denen der klassischen Testtheorie (KTT), sind aber vor dem Hintergrund der Lernverlaufsdagnostik anders gewichtet und führen in der Umsetzung zu Schwierigkeiten (Wilbert & Linnemann, 2011). Deshalb werden sie in der aktuellen Forschungsliteratur als noch offene Forschungsfragen diskutiert. Klauer (2011) sowie Wilbert und Linnemann (2011) nennen folgende zentrale Anforderungen an Verfahren der Lernverlaufsdagnostik: Reliabilität und Änderungssensitivität, Validität, Eindimensionalität, gleiche Schwierigkeit der einzelnen Tests, Testfairness sowie die Frage, nach welchem testtheoretischen Ansatz, ein Verfahren der Lernverlaufsdagnostik konstruiert werden sollte. Vor dem Hintergrund des Ziels dieser Arbeit, ein Verfahren der Lernverlaufsdagnostik zu entwickeln, werden diese Anforderungen nachfolgend erläutert.

Reliabilität und Änderungssensitivität

Soll ein Verfahren Leistungsveränderungen dokumentieren, muss es reliabel und

veränderungssensitiv sein. Ein reliables Verfahren der Lernverlaufsdagnostik soll das zu erfassende Merkmal also möglichst genau messen, aber zugleich auch Veränderungen sensitiv erfassen. Diese Anforderung ist nicht mit der KTT vereinbar, denn die KTT geht bei Messwiederholungen von einer zeitlichen Stabilität der Merkmale aus und behandelt Veränderungen als Messfehler. Deshalb ist eine hohe Retestreliabilität mit der Veränderungsmessung unvereinbar, da hohe Werte der Retestreliabilität dafür stehen, dass keine oder identische Veränderungen stattgefunden haben. Veränderungssensitive Verfahren sollten daher nur mittelhohe Retestreliabilitäten, aber einen hohen Wert der Testhalbierungsreliabilität aufweisen. Zugleich sollten die Korrelationen zwischen den Verfahren über einen längeren Zeitraum abnehmend (Klauer, 2011; Wilbert & Linnemann, 2011).

Validität

Verfahren der Lernverlaufsdagnostik, die eine gleiche Leistung über einen längeren Zeitraum erfassen sollen, müssen zu den jeweiligen Messzeitpunkten das gleiche Merkmal messen, dieses also valide erfassen. Entsprechend dieser Anforderung sollten Verfahren der Lernverlaufsdagnostik nicht nur das erfassen, was aktuell im Unterricht vermittelt wird, denn die Inhalte und Anforderungen ändern sich innerhalb eines Schuljahres. Das Lehrziel ändert sich, es wird nicht mehr das Gleiche gemessen. Damit würde sich sodann auch die Validität der Verfahren verändern. Verfahren der Lernverlaufsdagnostik sollten daher nicht das messen, was aktuell gelernt wird, sondern die Leistungen erfassen, die am Ende eines Schuljahres beherrscht werden sollten. Somit wird zu jedem Messzeitpunkt das gleiche Merkmal valide erfasst (Klauer, 2011).

Eindimensionalität

Soll ein Verfahren Veränderungen einer Leistung erfassen, „muss das Ergebnis einer Messung auch nur auf diese eine Fähigkeit zurückzuführen sein“ (Wilbert & Linnemann, 2011). Die Forderung nach Eindimensionalität bedeutet also z. B., dass ein Verfahren für den Lernbereich Rechtschreiben nur Rechtschreibfähigkeiten, nicht aber Anstrengungsbereitschaft oder Testängstlichkeit erfasst (Wilbert & Linnemann, 2011).

Testschwierigkeit

Ein Verfahren der Lernverlaufsdagnostik besteht aus einzelnen Paralleltests, die die gleiche Leistung über einen längeren Zeitraum erfassen. Diese Paralleltests müssen gleich

schwer sein und die gleiche Trennschärfe haben, um wahre Leistungsveränderungen dokumentieren zu können. Würden die Schwierigkeiten des Verfahrens bei Messwiederholungen schwanken, könnten keine gültigen Aussagen über Leistungsveränderungen getroffen werden (Klauer, 2011; Strathmann & Klauer, 2010; Wilbert & Linnemann, 2011).

Testfairness

Ist ein Test fair, haben Personen mit gleich hoher Merkmalsausprägung, die gleiche Wahrscheinlichkeit einen bestimmten Test zu lösen. Ist dies aufgrund von Nationalität, Alter oder Geschlecht nicht so, ist der Test unfair (Wilbert & Linnemann, 2011).

Testtheorie

Für die Entwicklung und Evaluierung von Verfahren der Lernverlaufsdiagnostik schlägt Klauer (2011) auf der Grundlage des Konzepts der kriteriumsorientierten bzw. lehrzielorientierten Tests das *Itemsampling* und das *Binomialmodell* vor, Wilbert und Linnemann (2011) schlagen eine *Verknüpfung der KTT und der Probabilistischen Testtheorie* (PTT) vor. Das Konzept der kriteriumsorientierten bzw. lehrzielorientierten Tests zielt auf die Erfassung einer Leistung, die durch ein vorher festgelegtes Kriterium (Lehrziel) festgelegt ist (Fisseni, 2004). Für die Entwicklung von Verfahren der Lernverlaufsdiagnostik wird eine das Kriterium betreffende relativ homogene Grundmenge an Aufgaben generiert und für jeden Schüler eine eigene Aufgabenstichprobe per Zufall erzeugt. Dieses Vorgehen heißt Itemsampling. Verfahren, die mittels Itemsampling konstruiert werden, sind nach dem Binomialmodell auszuwerten (Klauer, 2011). „Das Binomialmodell . . . beruht auf der Grundannahme, dass alle Aufgaben, die das gleiche latente Merkmal anzielen, die gleiche Schwierigkeit haben. Die Binomialverteilung gibt dann für jeden möglichen Testwert (jede Testaufgabe) die Wahrscheinlichkeit seines Auftretens an“ (Ingenkamp & Lissmann, 2008, S. 128). Wilbert und Linnemann schlagen für die Analyse von Verfahren der Lernverlaufsdiagnostik eine vierstufige Vorgehensweise vor:

1. Itemanalyse (Trennschärfe, Itemschwierigkeit, Interne Konsistenz, Homogenität, Reliabilitätsschätzung) nach der KTT,
2. Prüfung der Eindimensionalität mittels konfirmatorischer Faktorenanalyse,
3. Analyse der Itemschwierigkeit mittels erweiterter Raschmodelle und
4. Analyse der Testfairness mittels logistischer Regressionsmodelle.

Insgesamt scheint eine gleichzeitige Anwendung der KTT und PTT für die Entwicklung und Evaluierung von Verfahren der Lernverlaufsdiagnostik angebracht, denn die KTT und PTT werden nicht als konkurrierende, sondern als ergänzende Modelle verstanden (Moosbrugger, 2007; Rost, 1996).

Die vorliegende Arbeit versucht die genannten offenen Forschungsfragen in der Entwicklung und Evaluierung der FE-RS 2 zu berücksichtigen, stützt sich aber maßgeblich auf das methodische Vorgehen der Arbeiten von Diehl (2008) und Knopp (2010). Hauptaugenmerk und Ziel der Arbeit ist, „den Lehrkräften Instrumente an die Hand zu geben, die es gestatten, auf einer einheitlichen Skala den Lernzuwachs während des Schuljahres regelmäßig und mit wenig Aufwand zu messen . . . Die Möglichkeiten dafür zu schaffen[,] wäre eine lohnende Aufgabe für Forschung und Entwicklung“ (Strathmann, Klauer & Greisbach, 2010, S. 76). Dieser Aufgabe widmet sich die vorliegende Arbeit und orientiert sich dabei an folgender Fragestellung: *Ist die FE-RS 2 ein valides, reliables und veränderungssensitives Verfahren, das die Lernentwicklung und den Lernverlauf der Schüler im Lernbereich Rechtschreiben der 2. Klassenstufe erfassen kann?*

1.3 Vorgehen und Aufbau der Arbeit

Die vorliegende Arbeit orientiert sich am Aufbau empirischer Arbeiten. Die theoretischen Überlegungen zu zwei diagnostischen Ansätzen und zum Rechtschreiberwerb führen zu Schlussfolgerungen für die Entwicklung der FE-RS 2 und zur Hypothesenbildung. Die anschließende empirische Untersuchung dient der Hypothesenprüfung.

Kapitel 2 erläutert zwei Ansätze zur Diagnostik und Prävention von Lernschwierigkeiten, den Aptitude-Treatment Interaction-Ansatz und den Response to Intervention-Ansatz. Ausgehend von einer hohen Bedeutung und Notwendigkeit evaluativer Diagnostik und schulischer Prävention werden die Entstehung und die Ziele der genannten Ansätze dargestellt sowie die Möglichkeiten zur Umsetzung beider Ansätze eingeschätzt. Das Kapitel schließt mit Schlussfolgerungen für die Entwicklung der FE-RS 2.

Kapitel 3 setzt sich mit den theoretischen Grundlagen des Rechtschreiberwerbs und den curricularen Vorgaben des Lernbereichs Rechtschreiben für die 2. Klassenstufe

auseinander. Vorab werden neben der Definition wesentlicher Begriffe, wie z. B. Schriftsprache und Rechtschreibung, die linguistischen Grundlagen des Schriftspracherwerbs erläutert. Es wird ein Überblick über den Aufbau des deutschen Sprach- und Schriftsystems sowie die Prinzipien und Regeln der Rechtschreibung gegeben, um die Besonderheiten des Lerngegenstandes Schriftsprache zu verdeutlichen. Anschließend werden Modelle der Rechtschreibentwicklung von Frith (1985, 1986), K. B. Günther (1986), Scheerer-Neumann (2008), May (2002) und Löffler und Meyer-Schepers (2008) vorgestellt. Daran schließt sich die Betrachtung von Modellen des Rechtschreibprozesses an: Ausgehend vom Modell von D. P. Simon und H. A. Simon (1973) werden Zwei-Wege-Modelle und Netzwerkmodelle erläutert. Im Anschluss daran werden die Anforderungen und Inhalte des Lernbereichs Rechtschreiben in den Bildungsstandards im Fach Deutsch für den Primarbereich (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK], 2005), in dem Rahmenplan Grundschule Deutsch von Brandenburg, Berlin, Bremen und Mecklenburg-Vorpommern (Ministerium für Bildung, Jugend und Sport des Landes Brandenburg [MBJS], Senatsverwaltung für Bildung, Jugend und Sport Berlin [SBJS], Senator für Bildung und Wissenschaft Bremen [SBW] & Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern [MBWK], 2004) und in den Sprachbüchern aus dem Schulbuchkatalog 2009/2010 für das Unterrichtsfach Deutsch der Jahrgangsstufe 2 in Grundschulen in Mecklenburg-Vorpommern (MBWK, 2009a) dargelegt. Das Kapitel schließt mit Schlussfolgerungen für die Entwicklung der FE-RS 2.

Kapitel 4 skizziert die Methodik der durchgeführten Untersuchung. Der Untersuchungsplan, die in der Studie verwendeten Untersuchungsinstrumente, die Stichprobenkonstruktion und die Untersuchungsdurchführung werden beschrieben und die abgeleiteten Forschungsfragen und Hypothesen sowie die verwendeten statistischen Analyseverfahren dargelegt.

Kapitel 5 umfasst die Ergebnisdarstellung zur empirischen Güte der FE-RS 2.

Kapitel 6 dient der Zusammenfassung und Diskussion der Ergebnisse. Zudem werden die Grenzen der Untersuchung deutlich gemacht und es folgt ein Vergleich der FE-RS 2 mit bestehenden Verfahren der Rechtschreibdiagnostik.

2 Diagnostik und Prävention von Lernschwierigkeiten im Aptitude-Treatment Interaction- und Response to Intervention-Ansatz

Vor dem Hintergrund der Notwendigkeit evaluativer Diagnostik und schulischer Prävention werden in diesem Kapitel zwei Ansätze für (sonder-) pädagogisches Handeln aufgezeigt. Ausgehend von Begriffsbestimmungen werden beide Ansätze hinsichtlich ihrer Entstehung, ihrer Ziele und ihren Möglichkeiten zur Umsetzung schulischer Prävention näher erläutert. Anschließend werden Schlussfolgerungen für die Entwicklung der FE-RS 2 gezogen.

2.1 Zur Bedeutung der evaluativen Diagnostik

„Sonderpädagoginnen und -pädagogen testen nicht, sondern betreiben Förderdiagnostik – so ein verbreitetes Credo. Förderdiagnostik wird dabei als besonders gute Diagnostik verstanden, als Diagnostik, welche nicht selektioniert, sondern ganzheitlich ist, sich am Individuum ausrichtet und auf Förderung hin ausgerichtet ist“ (Moser Opitz, 2006, S. 5). *Förderdiagnostik* impliziert nur Positives; man diagnostiziert, um zu fördern. Sie grenzt sich von der sogenannten *Selektionsdiagnostik* und der damit verbundenen Ausgrenzung in Förderschulen ab. Allerdings wurde der Begriff der Selektion in der Diskussion um Selektions- versus Förderdiagnostik zum Teil unzutreffend definiert, denn Selektion ist, neutral betrachtet, keine Entscheidung oder Maßnahme, die eine Person schädigt. Im Gegenteil, Selektion hat das Ziel, eine Passung zwischen Person und Maßnahme herzustellen. Eine Maßnahme in einem segregierenden Schulsystem kann auch die Auswahl eines Förderorts sein. Das Problem *Selektionsdiagnostik* entsteht erst dadurch, dass mit der Überweisung an die Förderschule die versprochene Förderung ausbleibt (Schlee, 2008; Schöler, 2008). Die Aussage, mit der Förderdiagnostik hätten angemessene Förderziele und Fördermaßnahmen Eingang in die sonderpädagogische Praxis gefunden, ist anzuzweifeln. Eine Förderdiagnostik, die diagnostische Ist-Werte erhebt, beruht auf einem „naturalistischen Fehlschluss“ (Schlee, 2008, S. 124). Sie geht davon aus, dass Diagnostik die Voraussetzung für sonderpädagogisches Handeln ist und dass aus Ist-Werten Soll-Werte und methodische Handlungshinweise abzuleiten sind. Die diagnostischen Daten zeigen jedoch nur an, was ist und was nicht ist, nicht aber, was sein

sollte und was nicht sein sollte. Förderziele und Fördermaßnahmen lassen sich aber nur auf Grundlage theoretischer Konzepte und Handlungsansätze und nicht allein von Diagnostik ableiten. Mit dem Begriff der Förderdiagnostik sind letztlich Wunschvorstellungen und Mythen verbunden, die nicht einlösbar sind (Schlee, 2007, 2008).

Trotz der Kritik an der Förderdiagnostik ist zu berücksichtigen, dass Diagnostik im (sonder-) pädagogischen Handeln eine wichtige Rolle einnimmt. Zukünftig sollte die Rolle der Diagnostik jedoch eine andere sein als bisher. Diagnostik sollte auf der Basis von Theorien erfolgen, genauso wie die sich anschließende Förderung, und die Nützlichkeit und Praktikabilität didaktischer und therapeutischer Konzepte überprüfen. Das bedeutet, dass Diagnostik nicht mehr wie bisher dem sonderpädagogischen Handeln *vorgeordnet* sein sollte, sondern den sonderpädagogischen Therapien und Förderkonzepten logisch und zeitlich *nachzuordnen* ist. Diagnostik wäre damit nicht mehr ein Entscheidungsinstrument für sonderpädagogisches Handeln, sondern ein theoriebasiertes Verbesserungsinstrument von sonderpädagogischem Handeln. Mit dieser Umkehrung bekäme die Diagnostik eine neue, *evaluative Funktion* (Schlee, 2004, 2007). „Wenn man die Diagnostik in diesem Sinne einsetzt, dann erfüllt sie eine eminent wichtige Aufgabe, weil dann mit ihrer Hilfe permanent an der Verbesserung von didaktischen und therapeutischen Handlungsempfehlungen gearbeitet werden kann“ (Schlee, 2007, S. 64). Ziel der evaluativen Diagnostik wäre dann, den Erfolg sonderpädagogischen Handelns zu erfassen und eine bestmögliche Förderung zu konzipieren. Eine solche Diagnostik setzt im schulischen Kontext valide Entwicklungsmodelle von schulisch geformten Fähigkeiten voraus.

Die evaluative Diagnostik kann entsprechend der gebräuchlichen Einteilung von Evaluationen nach dem Zeitpunkt der Datenerhebung *summativ* oder *formativ* sein. Die summative Evaluation ist die bekannteste Evaluationsform im pädagogischen Kontext. Sie dient der Überprüfung der Lernzielerreichung. Die summative Evaluation wird am Ende von Maßnahmen oder Unterrichtseinheiten durchgeführt und überprüft in Form von Klassenarbeiten und Prüfungen, ob diese erfolgreich waren, angestrebte Förderziele erreicht wurden und was Schüler gelernt haben. Die formative Evaluation wird in der pädagogischen Praxis weit weniger angewendet. Sie dient nicht der zusammenfassenden Überprüfung von Leistungen, sondern der verbesserten Gestaltung von Lernprozessen. Formative Evaluationen werden fortlaufend in kurzen Zeitabständen im Unterricht

durchgeführt, überprüfen pädagogisches Handeln auf Stärken und Schwächen sowie den Lernverlauf des Schülers und dienen der Reflexion des Unterrichts (Goethe-Institut Inter Nationes, 2001; Woolfolk, 2008). Auch wenn beide Evaluationsformen ihre Berechtigung im Unterricht haben, sind vor dem Hintergrund von Prävention und individueller Förderung vor allem formative Instrumente, die Lernprozesse und Lernverläufe ökonomisch und differenziert erfassen, notwendig. Sind diese formativen Instrumente theoriegeleitet, reliabel und valide, stellen sie fortlaufend eine enge Verbindung zwischen Diagnostik und pädagogischen Maßnahmen her und erlauben dadurch frühzeitiges präventives Handeln.

Im folgenden Abschnitt wird ausgehend von einer Begriffsbestimmung das Konzept und die Bedeutung der schulischen Prävention kurz erläutert.

2.2 Zur Bedeutung der schulischen Prävention

Obwohl die Vorbeugung von Krankheiten seit Jahrhunderten ein wichtiges Anliegen der Medizin ist, ist *Prävention* ein vergleichsweise junger Begriff. Erst in der 2. Hälfte des 20. Jahrhunderts wurde der Begriff in seiner heutigen Bedeutung geprägt und auf weitere wissenschaftliche Disziplinen übertragen. Ganz allgemein bedeutet Prävention Vorbeugung und Verhütung und umfasst Maßnahmen, um unerwünschte Entwicklungen oder Ergebnisse zu verhindern (Jogschies, 2008). Die bekannteste medizinische Klassifikation von Prävention ist die von Caplan (1964). Caplan (1964) unterscheidet Prävention nach dem Zeitpunkt der Präventionsmaßnahmen in *primäre*, *sekundäre* und *tertiäre Prävention*. Die primäre Prävention zielt darauf ab, durch Aufklärungsmaßnahmen und konkrete Angebote das Auftreten von Erkrankungen zu verhindern und die Anzahl von Neuerkrankungen zu verringern. Primärprävention findet somit vor Eintritt der Erkrankung statt. Die sekundäre Prävention zielt darauf ab, Erkrankungen frühzeitig zu erkennen und effektiv zu behandeln, um das Fortschreiten dieser zu verhindern. Die tertiäre Prävention zielt darauf ab, die vollständig ausgebrochene Erkrankung zu behandeln sowie Folgeschäden und Rückfallrisiken zu verhindern bzw. zu verringern.

Ebenso wie in der Medizin hat das Konzept der Prävention auch in der Pädagogik Bedeutung erlangt. Prävention ist zu einer pädagogischen Zielsetzung geworden. Sie zielt

darauf ab, die schulischen Lern- und Lebensbedingungen bestmöglich für alle Schüler zu gestalten. Lernförderliche Bedingungen sollen geschaffen und optimiert, belastende und gefährdende Bedingungen reduziert werden. *Schulische Prävention* ist demnach „ein Ensemble von Maßnahmen . . ., welche geeignet sind zu verhindern, dass sich bei Kindern und Jugendlichen, welche von Behinderung bedroht sind . . ., manifeste Lern- und Verhaltensprobleme bilden“ (Kretschmann, 2000, S. 325). Gegenstandsbereiche schulischer Prävention sind neben Störungen, deren Ursachen außerhalb der Schule liegen, Lern- und Verhaltensschwierigkeiten, die durch inadäquaten schulischen Unterricht hervorgerufen oder verstärkt werden können (Kretschmann, 2000). Die Klassifikation nach Caplan (1964) ist auch auf den Bereich der schulischen Prävention übertragbar:

1. Primäre schulische Prävention beinhaltet eine kindgerechte Gestaltung der schulischen und außerschulischen Lern- und Lebensbedingungen. Primäre Prävention ist alles, was eine gute Schule und einen guten Unterricht ausmacht. Dazu zählen z. B. ein ansprechender und effizienter Unterricht, die Berücksichtigung unterschiedlicher Ausgangslagen und Binnendifferenzierung. Sonderpädagogischer Förderbedarf wird verhindert.
2. Sekundäre schulische Prävention umfasst die Früherkennung von sich anbahnenden Lern- und Verhaltensschwierigkeiten sowie schulische und außerschulische Unterstützungsangebote für Kinder, die von Schwierigkeiten und Schulversagen bedroht sind. Dazu zählen schulorganisatorische und pädagogische Konzepte, die geeignet sind, Lern- und Verhaltensschwierigkeiten vorzubeugen und einer Aussonderung von Schülern mit Schwierigkeiten entgegenwirken. Umfassende Lernrückstände und sonderpädagogischer Förderbedarf werden verhindert und die Teilnahme am regulären Unterricht ermöglicht.
3. Tertiäre schulische Prävention beinhaltet Förderung, Therapie sowie schulische und außerschulische Unterstützungsangebote bei manifesten Störungen. Der Entwicklung von möglichen externalisierenden und internalisierenden Verhaltensstörungen wird vorgebeugt. Die Förderung von Schülern mit sonderpädagogischem Förderbedarf erfolgt z. B. über die Vermittlung von Arbeitstechniken und die Förderung der emotionalen Stabilität (Jogschies, 2008; Kretschmann, 2000, 2003).

Die angeführten Überlegungen zur schulischen Prävention erscheinen sinnvoll und logisch. Auf Störungen wird nicht erst reagiert, sondern es werden Maßnahmen getroffen, damit diese nicht auftreten. Auch wenn in Wissenschaft, Bildungspolitik und Schulpraxis

schulische Prävention weitestgehend anerkannt ist (Hartmann & C. Müller, 2009), ist die Sonderpädagogik in Deutschland „jedoch traditionell remedial ausgerichtet, das heisst [sic] Schulkinder werden im Sinne eines ‚Wait-to-fail‘-Modells (Vaughn & Fuchs, 2003) erst dann sonderpädagogischen Interventionen zugeführt, nachdem ihre Lerndefizite manifest (verfestigt) und (spät) diagnostiziert sind“ (Hartmann & C. Müller, 2009, S. 25). Zudem ist das förderpädagogische Denken und Handeln nach wie vor durch den *Aptitude-Treatment Interaction-Ansatz* (ATI-Ansatz), vor allem das remediale Vorgehen, bestimmt (Hartmann & C. Müller, 2009). Das remediale Vorgehen beugt Lernschwierigkeiten aber nicht vor, sondern stellt Platzierungsstrategien und entsprechende Fördermaßnahmen erst bereit, wenn Lerndefizite und Förderbedarf bestehen. Es entspricht der tertiären schulischen Prävention, „das Kind ist dann bereits in den Brunnen gefallen“. Des Weiteren sind mit dem ATI-Ansatz theoretische, methodische und praktische Probleme verbunden. Um die Möglichkeiten schulischer Prävention zu nutzen, das heißt vor allem den Anteil von Schülern mit sonderpädagogischem Förderbedarf zu reduzieren, bedarf es neuer, präventiver und diagnostischer Vorgehensweisen. Möglicherweise bietet sich der aus den USA stammende und dort viel praktizierte Präventionsansatz *Response to Intervention* (RTI) an.

Im folgenden Abschnitt wird der ATI-Ansatz als bislang dominierender Ansatz der deutschen Förderpädagogik und -diagnostik erläutert. In Abschnitt 2.4 wird der RTI-Ansatz als Alternative für schulische Prävention und Diagnostik erläutert und anschließend mit dem ATI-Ansatz verglichen.

2.3 Aptitude-Treatment Interaction

Auf der Jahrestagung der American Psychological Association 1957 bewertete Cronbach die Trennung zwischen experimentellen und korrelativen Ansätzen der Psychologie als Hindernis für die psychologische Forschung. Um die Forschung voranzutreiben, forderte er die Verschmelzung beider Ansätze und eine Disziplin, die Persönlichkeitsmerkmale und Treatments gleichzeitig als unabhängige Variablen betrachtet. Cronbach (1957) prägte damit den Begriff *Aptitude-Treatment Interaction* und begründete die Notwendigkeit des ATI-Ansatzes wie folgt:

„We can expect some attributes of persons to have strong interactions with treatment variables. These attributes have far greater practical importance than the attributes which have little or no interaction. . . . We require a measure of aptitude which predicts who will learn better from one curriculum than from the other; but this aptitude remains to be discovered. Ultimately we should *design* treatments, not to fit the average person, but to fit groups of students with particular aptitude patterns. Conversely, we should seek out the aptitudes which correspond to (interact with) modifiable aspects of the treatment“ (S. 680f.).

Das ATI-Konzept ist ein Ansatz zur Individualisierung und Differenzierung des Unterrichts. Der ATI-Ansatz fordert die Entwicklung alternativer Unterrichtsmethoden, die sich den individuellen Unterschieden der Schüler anpassen.

2.3.1 Begriffsbestimmungen

Für die Auseinandersetzung mit dem ATI-Ansatz ist die Bestimmung von Grundbegriffen unerlässlich. Deshalb werden dem ATI-Ansatz zugrunde liegende Begriffe definiert.

Aptitudes sind stabile Persönlichkeitsmerkmale. Das Konzept der Aptitudes ist mit Begriffen wie Potenzial, Befähigung und Bereitschaft verknüpft. Aptitudes beschreiben individuelle Unterschiede im Lernen und Verhalten in spezifischen Situationen. Das Konzept der Aptitudes betont die charakteristische Wechselwirkung, das heißt die gegenseitige Passung oder Fehlpassung, zwischen einer Person und einer Situation. Daher rufen Aptitudes in unterschiedlicher Ausprägung, wie z. B. Intelligenzleistungen und Einstellungen, individuell unterschiedliche Person-Situation-Interaktionen hervor (Snow, 1991a, 1991b).

Der Begriff *Treatment* ist sehr allgemein. Im Wesentlichen sind Treatments komplexe Situationen, Umgebungen und Behandlungen. Sie umfassen variable Komponenten der Situation, z. B. klinische vs. private Umgebung, Gruppe vs. individual, und der in diesem Treatment agierenden Person, z. B. Alter und Geschlecht (Snow, 1991a). Treatments sind wie Aptitudes relationale Konzepte. Für Snow besteht die Bedeutung des relationalen Aspekts darin, dass Personen bis zu einem gewissen Grad versuchen, Treatments entsprechend ihrer Aptitudes zu gestalten und anzupassen.

Outcomes beschreiben Resultate und Ergebnisse von Treatments und Aptitudes. Gleiche

Treatments führen bei Personen mit Aptitudes unterschiedlichen Ausmaßes zu unterschiedlichen Outcomes (Caspi & Bell, 2004).

Interactions sind Wechselwirkungen zwischen Aptitudes und Treatments. Nach statistischer Definition liegt eine Interaktion vor, wenn zwei oder mehr unabhängige Variablen (Ausmaße des Aptitudes, Treatments) eine dritte abhängige Variable (Outcome) beeinflussen. Ein Aptitude wäre z. B. das Geschlecht (weiblich: A_1 /männlich: A_2), ein Treatment z. B. die Unterrichtsmethode (Frontalunterricht: B_1 /Offener Unterricht: B_2) und ein Outcome (y) wäre dann z. B. der Lernerfolg. In der ATI-Forschung wird die Wechselwirkung von Aptitudes und Treatments auf die Outcomes untersucht (R. Schwarzer & Steinhagen, 1975; Snow, 1991a). Lindquist (1953) unterscheidet zwischen *ordinaler* und *disordinaler* Interaktion. R. Schwarzer und Steinhagen (1975) ergänzen die beiden Interaktionsarten um die *Nicht-Interaktion*. Abbildung 1 stellt die beiden Interaktionsarten und die Nicht-Interaktion in Anlehnung an R. Schwarzer und Steinhagen (1975) graphisch dar.

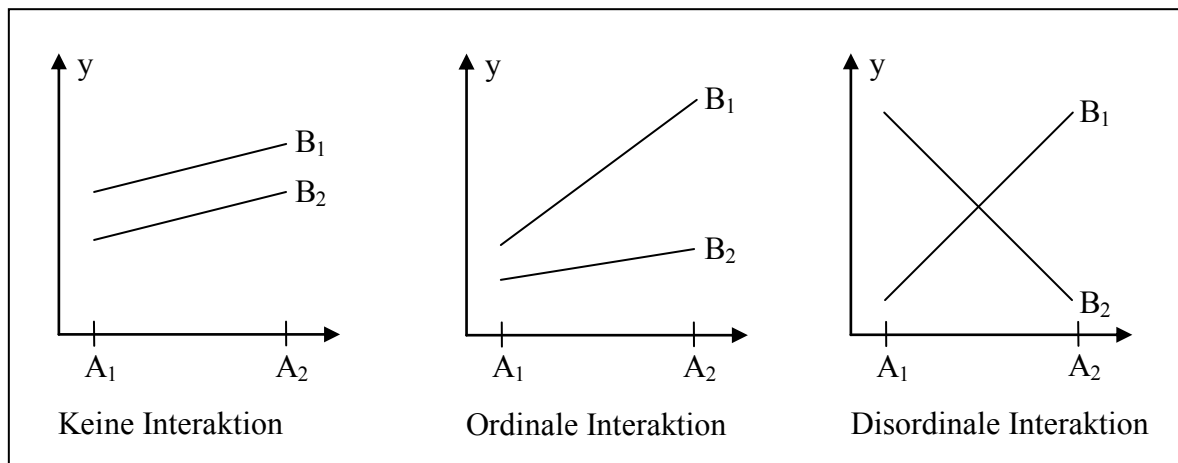


Abbildung 1: Die Interaktionsarten (in Anlehnung an R. Schwarzer & Steinhagen, 1975, S. 16)

Anmerkungen: A_1 und A_2 = Aptitude, B_1 und B_2 = Treatment, y = Outcome

Eine Nicht-Interaktion liegt vor, wenn Aptitude und Treatment nicht miteinander interagieren und jede unabhängige Variable unabhängig von der anderen auf das Outcome wirkt. Ein hypothetisches Beispiel; der Frontalunterricht ist dem Offenen Unterricht und die Jungen sind den Mädchen überlegen. Die beiden Graphen verlaufen daher in der graphischen Darstellung parallel zueinander. Eine ordinale Interaktion liegt vor, wenn Aptitude und Treatment miteinander interagieren und ein Treatment dem anderen Treatment überlegen ist. Ein Beispiel; die Jungen profitieren stärker von der Überlegenheit

des Frontalunterrichts als die Mädchen. Die beiden Graphen verlaufen daher in der graphischen Darstellung nicht parallel, kreuzen sich aber auch nicht. Eine disordinale Interaktion liegt vor, wenn sich die Wechselwirkung zwischen Aptitude und Treatment ändert und beide Treatments durchschnittlich gleich effektiv sind. Ein Beispiel; der Frontalunterricht ist für die Jungen gut und für die Mädchen weniger gut geeignet, für den Offenen Unterricht ist es umgekehrt. Die beiden Graphen kreuzen sich daher in der graphischen Darstellung. Die disordinale Interaktion ist der ATI-Effekt, den ATI-Studien ausfindig machen wollen (R. Schwarzer & Steinhagen, 1975).

Typische ATI-Studien untersuchen die Wechselwirkungen zwischen Unterrichtsmethoden und Schülermerkmalen (Snow, 1991a). Eine ATI liegt vor, wenn ein Schüler aufgrund eines bestimmten Aptitudes besser lernt, wenn er nach einer bestimmten, für ihn geeigneten Unterrichtsmethode unterrichtet wird, während ein anderer Schüler, der das betreffende Aptitude in geringerem Maße besitzt, wiederum nach einer anderen, für ihn geeigneten Unterrichtsmethode besser lernt (Gage & Berliner, 1996).

2.3.2 Entstehung des Ansatzes

Die Grundidee des ATI-Ansatzes, Unterricht durch Berücksichtigung individueller Unterschiede zu verbessern (Salomon, 1972, 1975), findet sich bereits in vorchristlichen Schriften (Corno et al., 2002; Snow, 1982). Die wissenschaftliche Definition von ATI als einheitliches Rahmenmodell der psychologischen Forschung stammt von Cronbach aus dem Jahr 1957. Das Interesse am ATI-Ansatz entwickelte sich in den folgenden Jahren zu einer ausgeprägten ATI-Forschung (Snow, 1991a). Die ersten Forschungsanstrengungen zwischen 1960 und 1970 waren aber zum Teil erfolglos und führten zu nicht zufriedenstellenden Ergebnissen (Cronbach, 1975a; Cronbach & Snow, 1977; Snow, 1989). Dennoch wurde die erste Dekade der ATI-Forschung positiv beurteilt, denn aus den bisherigen Forschungsanstrengungen konnten folgende Schlussfolgerungen für die zukünftige ATI-Forschung gezogen werden (Cronbach & Snow; Snow, 1989):

1. ATIs existieren,
2. konventionelles theoretisches Denken und statistische Methoden werden komplexen Interaktionen nicht gerecht,
3. aufgrund der Komplexität, Flexibilität und Vielfalt von Unterricht gibt es keine

allgemeingültigen ATIs,

4. „general ability“ (Snow, 1989, S. 22) als allgemeine Lern- und Leistungsfähigkeit wirkt in Interaktionen häufiger als andere Aptitudes,
5. Art und Ausmaß der Strukturierung des Unterrichts wirkt in Interaktionen als Treatment am stärksten,
6. weitere Aptitudes wie Vorwissen, Ängstlichkeit und Motivation sind häufig replizierte Variablen in ATIs,
7. andere Aptitudes, wie Lehrerwahrnehmung durch Schüler, sind weniger bedeutsame Variablen in ATIs.

Neben theoretischen und methodischen Schwierigkeiten (Snow, 1989) war ein Problem der ATI-Forschung der ersten Dekade, dass Aptitudes und Outcomes als gegeben und veränderbare Aptitude-Outcome-Beziehungen als Funktionen von Treatments betrachtet wurden. Bisher wurde nicht untersucht, wie Aptitudes und Outcomes miteinander verbunden sind, sodass unterschiedliche Treatments zu veränderten Aptitude-Outcome-Beziehungen führen können (Snow, 1980). Aufgrund dessen gab es zwei große Aufgabenfelder für die zukünftige Forschung: Zum einen die Untersuchung des komplexen, realen schulischen und sozialen Umfelds, in dem ATIs stattfinden und zum anderen die Untersuchung des Lernprozesses, in dem Aptitudes und Treatments wirken. In den Folgejahren weitete sich die ATI-Forschung aus und Untersuchungen zum Zusammenhang zwischen Aptitudes und Lernerfolg wurden durchgeführt (Snow, 1989). Diese Untersuchungen belegten die Bedeutung des Zusammenhangs. „Aptitude process differences exist before and operate through, but are also produced by, instruction to account for individual differences in learning outcome“ (Snow, 1980, S. 29).

In der zweiten Dekade, 1975 bis 1985, wurde die Analyse der Aptitude Processes fortgeführt (Snow, 1980). Gleichzeitig wendeten sich, insbesondere in den 1980er Jahren, viele Psychologen von der ATI-Forschung aufgrund der theoretischen und methodischen Probleme ab (Corno et al., 2002). Snow (1980) bezeichnete die ATI-Forschung in der zweiten Dekade als „process-oriented research on individual differences in learning and cognition“ (S. 28). Mit dieser Phase war die Hoffnung verbunden, durch neuere Lern- und Entwicklungstheorien sowie Intelligenzkonzepte *neue* Aptitudes zu entdecken (Glaser, 1972, 1975). Eine neue Aptitude-Theorie konnte in der zweiten Dekade zwar nicht konstruiert werden, aber die ATI-Forschung belegte, dass nicht, wie früher angenommen, allein kognitive Aptitudes individuelle Unterschiede im Lernen hervorrufen, sondern sich

gute und weniger gute Lerner durch das Zusammenspiel kognitiver, motivational-volitionaler und affektiver Aptitudes unterscheiden (Snow, 1980, 1989).

In den 1980er und 1990er Jahren rückten die motivational-volitionalen und affektiven Aptitudes in den Mittelpunkt der ATI-Forschung (Corno et al., 2002). Eine ausgeprägte Forschung um das Phänomen ATI war aber nicht mehr erkennbar. Mit dem Tod von Snow 1997 schien die ATI-Forschung beendet. Durch die intensiven Debatten um die Qualität von Unterricht und Schule in Folge von zahlreichen Schulleistungsstudien kommt heute die Frage nach Leistungsunterschieden sowie der Wechselwirkung zwischen Unterricht und Schülermerkmalen erneut auf. Daraufhin wurden bereits einige Arbeiten, die auf dem ATI-Ansatz basieren, veröffentlicht (z. B. E. Beck et al., 2008; Reisinger, 2007).

Die ATI-Forschung entstand nicht aus dem Bruch und der Unzufriedenheit mit vorherrschenden Unterrichtsmethoden. Die ATI-Forschung entstand vielmehr aus allgemeinen psychologischen Überlegungen über statistische Methoden und über die Struktur menschlicher Fähigkeiten. Erst später wurden innerhalb der ATI-Forschung Probleme des Lehrens und Lernens sowie der Interaktion erörtert (Cronbach & Snow, 1969; Kallós, 1975). Die mit dem ATI-Ansatz verbundenen Ziele werden im nachfolgenden Abschnitt beschrieben.

2.3.3 Ziele und Funktionen des Ansatzes

Das unmittelbare Ziel der ATI-Forschung im Kontext von Unterricht ist, für ausgewählte Schülermerkmale die entsprechend passende Unterrichtsmethode zu finden, um definierte Ziele zu erreichen (Cronbach & Snow, 1969). Die Suche nach der besten Unterrichtsmethode für alle Schüler wird aufgegeben, da verschiedene Schüler nach verschiedenen Methoden unterrichtet werden sollten (Salomon, 1972, 1975; Kallós, 1975). Das mittelbare Ziel der ATI-Forschung ist, dass sich der ATI-Ansatz auf die Lehrerbildung auswirkt. Es wird angestrebt, dass Merkmale der Lehrerpersönlichkeit und des Unterrichtsstils sowohl als Aptitude, in Bezug auf die Lehrerausbildung, als auch als Treatment, welches das Lernen der Schüler beeinflusst, verstanden werden (Cronbach & Snow, 1969).

Die ATI-Forschung hat zwei Funktionen: die Verbesserung des Unterrichts und die Entwicklung von Erklärungsprinzipien für das, was Unterricht ausmacht (Salomon, 1972, 1975). Die ATI-Forschung berücksichtigt individuelle Unterschiede, denn sie basiert auf der Annahme, dass bestimmte Unterrichtsmethoden für manche Schüler förderlich oder nicht förderlich sind. Das Ziel des Ansatzes ist, dass Schüler mit unterschiedlichen Voraussetzungen, unter der Verwendung von auf ihren Aptitudes abgestimmten Unterrichtsmethoden, einen gleich hohen Lernerfolg erzielen. Eine Verbesserung des Unterrichts wird demzufolge angestrebt (Reisinger, 2007; Salomon, 1972, 1975). Die zweite Funktion der ATI-Forschung dient der Weiterentwicklung erklärender Prinzipien für den Unterricht. „Die ATI-Forschung könnte die Entwicklung einer Unterrichtstheorie dadurch begünstigen, daß [sic] sie nach und nach eine Matrix von Lernsituationen und Charakteristika des Lernenden konstruiert“ (Salomon, 1975, S. 128).

Der ATI-Ansatz steht in der Tradition der Individualisierung von Unterricht und soll stabile Wechselwirkungen zwischen Aptitudes und Treatments nachweisen. Gruppierungen nach Aptitudes und entsprechenden Treatments sind die Folge. Demnach sollen sich nicht die Schüler dem Unterricht anpassen, sondern umgekehrt. Der ATI-Ansatz fordert die Berücksichtigung unterschiedlicher Schülermerkmale durch den Unterricht. Welche Modelle der Individualisierung und der Adaption des Unterrichts es im Rahmen des ATI-Ansatzes gibt, zeigt der nachfolgende Abschnitt auf.

2.3.4 Modelle der Adaption

Die adaptive Gestaltung von Lehr-Lernsituationen hat, wie bereits erwähnt, eine lange Tradition. Dabei wird der Begriff *Adaption* in der Forschung unterschiedlich verwendet (Leutner, 1992). Leutner systematisiert das Konzept der Adaption und differenziert nach *Adaptionsrate*, *Adaptionsmaßnahme* und *Adaptionszweck*.

Die Adaptionsrate beschreibt den zeitlichen Abstand zwischen Anpassungen (Leutner, 1992). Cronbach (1967, 1975b) unterscheidet zwischen einer *Mikro-Adaption* und einer *Makro-Adaption*. Mikro-Adaptionen bezeichnen Anpassungsentscheidungen, die im Abstand von Sekunden und Minuten sowie während des Unterrichts getroffen werden. Sie sind im Hinblick auf ATIs genauer, aber nicht besser, weil sie sich nur auf wenige

Informationen stützen. Mikro-Adaptionen sind Gegenstand dynamisch, adaptiver Lehrsysteme. Makro-Adaptionen bezeichnen Anpassungsentscheidungen, die im Abstand von Wochen und Monaten sowie vor Beginn des Unterrichts getroffen werden. Sie können theoriegeleitet konstruiert und empirisch überprüft werden. Makro-Adaptionen sind Gegenstand der ATI-Forschung (Cronbach, 1967, 1975b; Leutner, 1992).

Die Adaptionsmaßnahme beschreibt, welche Maßnahmen zur Anpassung an den einzelnen Schüler vorgenommen werden. Zu den Adaptionsmaßnahmen zählen die Anpassung durch Veränderung der *Lehrzeit*, der *Lehrziele* und der *Unterrichtsmethoden*. Lehrzeiten, Lehrziele und Unterrichtsmethoden sind elementare Adaptionsmittel, die einzeln, aber auch in Kombination auftreten können. Durch die Kombination ergeben sich acht Adaptionsmaßnahmen. Sie können festgelegt/fixiert oder variabel/angepasst sein (Leutner, 1992). Die acht Adaptionsmaßnahmen sind in Tabelle 1 in Anlehnung an Leutner (1992) dargestellt.

Tabelle 1: Adaptionsmaßnahmen (in Anlehnung an Leutner, 1992, S. 10)

Adaptionsmittel			Adaptionsmaßnahme (Beispiel)
Ziel	Methode	Zeit	
fixiert	fixiert	fixiert	Fortschreitende Auslese
		variabel	Mastery Learning bzgl. Lehrziel
	angepasst	fixiert	Innere (Unterrichts-) Differenzierung
		variabel	Förderunterricht/ Nachhilfeunterricht
angepasst	fixiert	fixiert	Notengebung beim herkömmlichen Unterricht
		variabel	Mastery Learning bzgl. Note
	angepasst (an Ziel)	fixiert	Äußere (Schul-) Differenzierung
		angepasst (an Methode)	Gegliedertes Schulsystem

Cronbach (1967, 1975b) unterscheidet fünf Adaptionsmaßnahmen. Hinsichtlich der Adaption der Lehrziele differenziert Cronbach (1967, 1975b) kaum, denn es sollten „bestimmte Lernziele von jedem erreicht werden . . . , infolgedessen jeder die Schule so lange besuchen sollte, bis er sie erreicht hat“ (Cronbach, 1975b, S. 44). Cronbach (1967, 1975b) unterscheidet folgende fünf Adaptionsmaßnahmen voneinander, wobei er der fünften Maßnahme im Rahmen des ATI-Ansatzes größte Bedeutung beimisst:

1. Variation der Schulzeit durch fortschreitende Auslese von Schülern,
2. Training einer Fertigkeit oder Thematik bis zum Lehrziel bei Variation der Schulzeit,
3. für jeden Schüler Bestimmung seiner voraussichtlichen Rolle als Erwachsener und Angebot eines Curriculums, das ihn auf die Rolle vorbereitet,
4. zusätzlicher Unterricht (Förderunterricht), der mit dem normalen Unterricht verzahnt ist und
5. verschiedene Methoden für verschiedene Schüler.

Der Adaptionzweck unterscheidet Adaptionmaßnahmen nach ihrem Zweck (Leutner, 1992). Salomon (1972, 1975) beschreibt drei Arten von Adaptionzwecken, sogenannte *heuristische Modelle*: das *Fördermodell (remedialer Ansatz)*, das *Kompensationsmodell* und das *Präferenzmodell*.

Das Fördermodell wird eingesetzt, um individuelle Lerndefizite durch Fördermaßnahmen zu überwinden. Es ist das am häufigsten praktizierte Modell. Ziel ist es, durch erhöhten Zeitaufwand und eine hohe Anzahl von fördernden Einheiten bisher nicht erreichte Lernziele zu erarbeiten. Entsprechende Schülermerkmale werden durch aufgabenspezifische Leistungstests ermittelt, denn von aufgabenspezifischen Fähigkeiten wird die Interaktion mit Treatments erwartet. Aufgabenspezifische Fähigkeiten werden in Lernziele transformiert und eingeübt. Das Vorgehen des Fördermodells ist hilfreich für Schüler, die bestimmte Lernziele nicht beherrschen. Für Schüler, die die notwendigen aufgabenspezifischen Fähigkeiten beherrschen, ist das Fördermodell hinderlich (Salomon, 1972, 1975).

Das Kompensationsmodell wird eingesetzt, um individuelle Lerndefizite durch geeignete Maßnahmen auszugleichen oder zu umgehen. Ziel ist es, durch Methoden, die bestimmte leistungsmindernde Persönlichkeitsmerkmale abbauen, nachteilige Effekte dieser Merkmale zu verhindern. Entsprechende Schülermerkmale werden durch Persönlichkeitstests, Intelligenz-, Angst- und Leistungstests ermittelt. Das Vorgehen des Kompensationsmodells ist hilfreich für Schüler, die nicht leistungsfähig genug sind, um sich selbstständig solche Methoden anzueignen. Für leistungsfähige Schüler, die sich selbstständig Methoden aneignen können, ist das Kompensationsmodell hinderlich (Hasebrook, 2006; Salomon, 1972, 1975).

Das Präferenzmodell wird eingesetzt, um individuell gut ausgeprägte Fähigkeiten durch die Anpassung des Lerninhalts, der Lernstruktur und der Präsentationsmodi zu nutzen. Ziel ist es, durch die Nutzung individueller Stärken Lerndefizite zu kompensieren. Entsprechende Schülermerkmale, wie Lernstil oder Art der Informationsverarbeitung, werden durch jegliche Art psychodiagnostischer Tests ermittelt. Das Vorgehen des Präferenzmodells ist hilfreich für Schüler, deren Fähigkeiten und Stärken, z. B. dass sie visuell orientiert sind, im Treatment enthalten sind. Für Schüler, deren ausgeprägte Fähigkeiten nicht im Treatment gefordert sind, ist das Präferenzmodell hinderlich (Hasebrook, 2006; Leutner, 1992; Salomon, 1972, 1975).

Die Adaptionrate und Adaptionsmaßnahme beziehen sich auf unterrichtstechnische Aspekte. Sie beschreiben, wie Unterricht realisiert wird, wenn das Ziel der Adaption festgelegt ist. Der Adaptionzweck legt das Ziel und die Funktion einer Adaptionsmaßnahme fest, wobei Unterrichtsmethoden selten nur eine definierte Funktion haben.

2.3.5 Kritik des ATI-Ansatzes

Der ATI-Ansatz hebt die Veränderbarkeit von Unterrichtsmethoden hervor und betont, Unterricht den individuellen Unterschieden der Schüler anzupassen. Der ATI-Ansatz ist damit eine Erinnerung daran, dass Schüler mit unterschiedlichen Fähigkeiten am besten nach verschiedenen Methoden lernen. Eine Methode allein garantiert nicht gleich hohe Lernerfolge für alle Schüler. Auch erzielen leistungsstarke Schüler nicht immer unabhängig von der Unterrichtsmethode bessere Lernergebnisse als leistungsschwache Schüler (Gage & Berliner, 1996; Kallós, 1975).

Trotz der mit dem ATI-Ansatz verbundenen Hoffnungen ist es in der heutigen Forschung wenig präsent. Die ATI-Forschung ist seit ihrer Entstehung von theoretischen, methodischen und praktischen Schwierigkeiten geprägt. Theoretische Schwierigkeiten ergaben sich z. B. durch die große Ähnlichkeit der Treatments untereinander und der Globalität der Aptitudes, z. B. Begabung. Zudem fehlte der ATI-Forschung eine theoretische Grundlage zur Auswahl von Prädiktoren, Lernbedingungen und deren Interaktion (Garten, 1980). „Es wird vermutet, daß [*sic*] ein grundlegender Widerspruch

zwischen dem statischen ATI-Ansatz, der über eine gewisse Zeit eine konstant günstige Zuordnung postuliert, und der Dynamik dabei ablaufender psychischer Prozesse besteht“ (Garten, 1980, S. 15). Methodische Schwierigkeiten ergaben sich z. B. dadurch, dass sich valide und wirkungsvolle Interaktionen in Experimenten, nicht aber in Unterrichtssituationen nachweisen ließen (Walter, 2008a; Wember, 2001). „Die empirische Validierung solcher Interaktionen zwischen Individuum und Umgebung hat sich . . . als ausgesprochen schwierig herausgestellt, weil es sich nicht um einfache und stabile, sondern um multipel moderierte, aufgaben- und situationsspezifische Interaktionen handelt“ (Wember, 2001, S. 179). Dennoch ist davon auszugehen, dass es ATIs gibt. Weitere Schwierigkeiten ergaben sich in der Schulpraxis. Der ATI-Ansatz fordert einen hohen Zeitaufwand für die jeweils individuelle Planung der Unterrichtsinhalte und Unterrichtsmethoden. Zudem fehlen für die hohen Anforderungen an die Lehrkräfte, z. B. die zusätzlichen Fördermaßnahmen, personelle Hilfen im Unterricht. Wollte eine Lehrkraft allen unterschiedlichen Lernvoraussetzungen der Schüler gerecht werden, könnte sie die Vielzahl von Differenzierungsmaßnahmen nicht bewältigen (Wember, 2001).

Der ATI-Ansatz bestimmt das derzeitige förderpädagogische Denken und Handeln, indem Gruppierungen nach Aptitudes und Treatments vorgenommen werden. Schüler mit niedrigen Lernvoraussetzungen, die Lehr- und Lernziele nicht erreichen, werden der Aptitude-Gruppe *lernbehindert* zugeordnet. Ihnen wird wiederum das Treatment *Förderschule* zugewiesen. Stabile Interaktionsbeziehungen, wie es der ATI-Ansatz fordert, können aber zwischen dem Aptitude *lernbehindert* und dem Treatment *Förderschule* im Gegensatz zum Treatment *allgemeine Schule* nicht nachgewiesen werden (Walter, 2008a). Die in der Schulpraxis gängigen Platzierungsstrategien sind daher nicht mehr als selektive Maßnahmen. Folglich sind förderpädagogisches Denken und Handeln nur durch scheinbare Adaption bestimmt. Wember (2001) führt dazu aus:

„Zwar wird mit der Überweisung an eine andere Schule dem Schulversager mittelbar ein Unterricht auf einem niedrigeren Anforderungsniveau zugewiesen, aber grundsätzlich gilt: Nicht die Unterrichtsinhalte und -methoden werden angepasst, sondern die Schülerinnen und Schüler werden bestimmten Niveaugruppen zugewiesen und haben sich dem Unterricht anzupassen. Die Lernvoraussetzungen der Lernenden und der Unterricht werden als Konstanten beibehalten, nicht als veränderliche Größen“ (S. 165).

Würde das Bildungssystem nach adaptiven Grundsätzen handeln, gäbe es eine Vielzahl einsetzbarer Adaptionsmaßnahmen, die sich individuell in Lehrziel, Lehrzeit und Methode

unterscheiden (Wember, 2001). In der Folge wären schulische Prävention und die Früherkennung von Lernschwierigkeiten möglich und das segregierende Schulsystem in Frage gestellt. Andererseits scheinen die bereits genannten Schwierigkeiten zu groß, um den ATI-Ansatz für adäquate Individualisierung und Differenzierung sowie Prävention anzuwenden. Ob der RTI-Ansatz die Schwierigkeiten des ATI-Ansatzes überwinden kann, legt der nachfolgende Abschnitt dar.

2.4 Response to Intervention (RTI)

Die in den USA erlassenen Gesetze „No Child Left Behind Act of 2001“ (NCLB) (Senate and House of Representatives of the United States of America in Congress, 2002) und „Individuals with Disabilities Education Improvement Act of 2004“ (IDEIA) (Senate and House of Representatives of the United States of America in Congress, 2004) heben die Bedeutung qualitativ hochwertigen Unterrichts und wissenschaftlich basierter Intervention hervor. Beide Gesetze betonen die Verantwortung der Schulen für die schulischen Leistungen der Schüler. Damit ist zugleich die Forderung nach systematischen und auf die Schüler abgestimmten pädagogischen Maßnahmen verbunden. Das IDEIA-Gesetz führt dazu aus, dass „a local educational agency may use a process that determines if the child responds to scientific, research-based intervention as a part of the evaluation procedures“ (§ 614, Absatz b6B). Der Begriff *process* verweist auf den Ansatz *Response to Intervention* (RTI) (U.S. Department of Education, 2006). Die gesetzliche Grundlage des RTI-Ansatzes ist damit das IDEIA-Gesetz. Die wissenschaftliche Begründung für den RTI-Ansatz findet sich in der US-amerikanischen Forschung der *special education* und der Frage nach der Identifizierung von Kindern mit *learning disabilities* (LD) (Fletcher, 2008).

2.4.1 Begriffsbestimmungen

Für die Auseinandersetzung mit dem RTI-Ansatz ist die Bestimmung von Grundbegriffen unerlässlich. Deshalb werden dem RTI-Ansatz zugrunde liegende Begriffe nachfolgend definiert.

R in RTI

Das *R* in RTI steht entweder für *response/responsiveness* oder für *resistance/non-responsiveness*. *Response/responsiveness* implizieren ein antwortbezogenes Vorgehen. Im Response to Intervention-Modell werden die Lehr- und Lernbedürfnisse eines Schülers aufgrund seiner positiven Reaktion auf eine Intervention, z. B. richtige Antworten und Lösungen, mit Hilfe von Screenings, Klassenarbeiten oder durch systematische Lernverlaufsdokumentationen diagnostiziert. Der Schwerpunkt dieses Vorgehens ist, zu ermitteln, welche Voraussetzungen ein Schüler hat und braucht, um vom Unterricht zu profitieren. Demzufolge ist das Response to Intervention-Modell eine Methode zur Ressourcenbelegung. Das bedeutet, dass die individuellen Ressourcen eines Schülers festgestellt werden, um diese für weitere Aufgaben einzusetzen und die Lehr- und Lernbedingungen entsprechend anzupassen. *Resistance/non-responsiveness* implizieren ein nicht-reaktionsbezogenes Vorgehen. Im Resistance to Intervention-Modell werden die Lehr- und Lernbedürfnisse eines Schülers aufgrund seiner Fehlreaktionen auf eine Intervention mit Hilfe von Screenings, Klassenarbeiten oder durch systematische Lernverlaufsdokumentationen diagnostiziert. Der Schwerpunkt dieses Vorgehens ist, zu ermitteln, welche Störungen und Beeinträchtigungen ein Schüler hat. Demnach ist das Resistance to Intervention-Modell eine defizitorientierte Methode. Das bedeutet, dass die individuellen Lerndefizite eines Schülers festgestellt werden, um die Lehr- und Lernbedingungen entsprechend anzupassen (Christ, Burns & Ysseldyke, 2005). In der vorliegenden Arbeit wird der Begriff *response* verwendet, da sich dieser von der defizitorientierten Sichtweise abgrenzt und im wissenschaftlichen Sprachgebrauch gängig ist.

I in RTI

Das *I* in RTI steht für *instruction* oder für *intervention*. *Instruction* impliziert allgemeine Maßnahmen, z. B. regulären Unterricht. Das Ziel des Response to Instruction-Modells ist die Evaluation der Effektivität allgemeiner Maßnahmen und die Identifizierung von Schülern, die intensivere Maßnahmen benötigen. Der englische Begriff *intervention* impliziert intensive, systematische und sonderpädagogische Maßnahmen sowie mehrmalige Evaluationen. Das Ziel des Response to Intervention-Modells ist die Evaluation der Effektivität intensiver Maßnahmen, um diese zu bewerten und gegebenenfalls anzupassen. Response to Instruction-Modelle und Response to Intervention-Modelle unterscheiden sich in der Häufigkeit und dem Ziel der Evaluation.

Allgemeine Maßnahmen sind für alle und intensive Maßnahmen nur für einige Schüler vorgesehen (Salvia & Ysseldyke, 2007). In der vorliegenden Arbeit wird der deutsche Begriff Intervention als Oberbegriff verwendet und falls erforderlich spezifiziert.

Die National Association of State Directors of Special Education (NASDSE, n. d.) definiert RTI als einen systematischen Prozess, der einen auf die Bedürfnisse der Schüler abgestimmten, qualitativ hochwertigen Unterricht (Instruction/Intervention) bereitstellt und der versucht, die Lernentwicklung und Lernleistung der Schüler einzuschätzen, um wichtige pädagogische Entscheidungen zu treffen. Kurz, RTI ist die Suche nach „*which children need what services, delivered with how much intensity*“ (Gresham, VanDerHeyden & Witt, 2005, S. 3).

2.4.2 Entstehung des Ansatzes

Die ersten Forschungsaktivitäten um den RTI-Ansatz liegen in den 1960er in den USA, aber erst in den letzten Jahrzehnten hat die RTI-Forschung bei Forschern und Praktikern an Beachtung gewonnen. Der RTI-Ansatz ist aus der Unzufriedenheit mit bestehenden diagnostischen Verfahren, insbesondere der Diskrepanzdefinition von Lernschwierigkeiten, entstanden (Bradley et al., 2007; Shores & Bender, 2007). Nach der Diskrepanzdefinition liegt eine Lernschwierigkeit vor, wenn bei durchschnittlicher Intelligenz die Schulleistung unterhalb der Altersnorm liegt. Gemäß der Definition gibt es zwei Arten von Schülern mit schulischen Minderleistungen; durchschnittlich intelligente und intelligenzschwache Schüler. Diese Einteilung ist nur sinnvoll, wenn beide Merkmale, Schulleistung und Intelligenz, valide und reliabel gemessen und die Cutoff-Grenzen nicht willkürlich gezogen werden. Sind diese Voraussetzungen nicht gegeben, kann es zu Fehleinschätzungen kommen. Neben den erwähnten Schwierigkeiten ist ebenso in Frage zu stellen, dass per definitionem intelligenzschwache Schüler keine Rechen- oder Lese-Rechtschreibschwäche aufweisen können, weil diese eine durchschnittliche Intelligenz voraussetzen. Die Definition von Lernschwierigkeiten nach der Diskrepanzdefinition ist in mehrfacher Weise problematisch (Klauer & Leutner, 2007; Walter, 2008a). Die Notwendigkeit veränderter Begriffsbestimmungen und Diagnosemöglichkeiten von Lernschwierigkeiten sowie die Notwendigkeit, sich von dem bisherigen Vorgehen des „wait to fail“ (Shores & Bender, 2007, S. 3) abzuwenden, ist offensichtlich.

Einige Autoren (z. B. Daly III, Martens, Barnett, Witt & Olson, 2007; Gresham, 2007; Gresham et al., 2005; Vaughn & L. S. Fuchs, 2003) zitieren die Studie des National Research Council (NRC) von Heller, Holtzman und Messick (1982) als theoretische Grundlage für den RTI-Ansatz. Ausgangspunkt der NRC-Studie ist, „to understand a child’s learning problems, one must assess not only intellectual functioning and other aspects of the individual outside the intellectual domain but also the contribution of the child’s educational environment to his or her performance in school“ (Heller et al., 1982, S. xi). Aufgrund dessen betonen Heller et al. (1982), dass valide Entscheidungen über Sonderbeschulungen auf folgenden drei Kriterien beruhen sollten:

1. alternative Strategien im regulären Unterricht,
2. valide Einschätzungen und Bewertungen und
3. Evaluierung der Qualität und Wirksamkeit von Sonderbeschulung.

In den folgenden Dekaden folgten zahlreiche Forschungsanstrengungen und Diskussionen zum RTI-Ansatz in den USA. Die Forschungsanstrengungen weiteten sich aus und führten zu neuen Forschungsschwerpunkten, z. B. RTI in der Vorschule und in der Sekundarstufe. Das U.S. Department of Education unterstützt den RTI-Ansatz stark. 2001 wurde die „Commission on Excellence in Special Education“ gegründet. Ziel war, durch die Abkehr von der Diskrepanzdefinition und dem Anwenden des RTI-Ansatzes die sonderpädagogische Förderung zu verbessern (U.S. Department of Education, 2002). Auch die neuen Richtlinien des 2004 erlassenen IDEIA-Gesetzes (Senate and House of Representatives of the United States of America in Congress, 2004) erlauben den Einsatz von RTI zur Identifizierung von Schülern mit Lernschwierigkeiten und sprechen sich gegen die Diskrepanzdefinition aus (Bradley et al., 2007). Durch die wissenschaftliche Fundierung und die Einführung staatlicher Initiativen und Gesetze gewinnt der RTI-Ansatz bei Lehrkräften und in den Schulen der USA an vermehrter Aufmerksamkeit. Bereits 44 U.S. Bundesstaaten berichten, dass sie den RTI-Ansatz in Erwägung ziehen, planen oder bereits durchführen (Hoover, Baca, Wexler-Love & Saenz, 2008).

Im deutschsprachigen Raum beginnen gegenwärtig die ersten wissenschaftlichen Auseinandersetzungen mit dem RTI-Ansatz. Der RTI-Ansatz wird als alternatives Präventions- und Diagnostikmodell untersucht und Möglichkeiten der Adaption werden aufgezeigt (Diehl & Hartke, 2007; Hartmann, 2008; Hartmann & C. Müller, 2009; Parmar, 2008; Walter, 2008a).

Der RTI-Ansatz ist aus der Unzufriedenheit mit der Diskrepanzdefinition zur Identifizierung von Schülern mit Lernschwierigkeiten entstanden. Die theoretische Grundlage des RTI-Ansatzes bildet die NRC-Studie (Heller et al., 1982). Die gesetzliche Grundlage des RTI-Ansatzes in den USA ist das IDEIA-Gesetz von 2004 (Senate and House of Representatives of the United States of America in Congress, 2004). Die mit dem RTI-Ansatz verbundenen Ziele werden in Abschnitt 2.4.3 beschrieben.

2.4.3 Ziele des Ansatzes

Das übergeordnete Ziel des RTI-Ansatzes ist die Prävention, insbesondere die Früherkennung, von Lernschwierigkeiten. Das bisherige diagnostische Vorgehen nach der Diskrepanzdefinition verhinderte das frühzeitige Eingreifen und Intervenieren bei Lernschwierigkeiten, da häufig gewartet werden musste, bis die schriftsprachlichen bzw. mathematischen Leistungen hinter denen der Intelligenz lagen. Zudem führte dieses Vorgehen oft zu falsch positiven und falsch negativen Ergebnissen. Vertreter des RTI-Ansatzes wollen *alle* Schüler durch den wiederholten und zeitlich engmaschigen Einsatz von Screenings, Lernerfolgskontrollen und Verfahren der Lernverlaufsmessung nach potenziellen Lern- und Verhaltensschwierigkeiten untersuchen. Damit ist das Ziel verbunden, den Unterricht entsprechend der Fähigkeiten der Schüler so zu gestalten, dass alle Schüler effektiv unterrichtet werden und dass dadurch das Auftreten von Lernschwierigkeiten verhindert wird. Außerdem intendiert dieses Vorgehen, Schüler mit Lern- und Verhaltensschwierigkeiten so früh wie möglich zu identifizieren, um zeitnah evidenzbasierte Fördermaßnahmen einleiten zu können. Die Früherkennung und Prävention von Lernschwierigkeiten verhindern außerdem schulisches Scheitern, Etikettierungsprozesse und selektive Maßnahmen. Es wird nicht gewartet bis „das Kind in den Brunnen gefallen ist“, sondern es werden vorab pädagogische Maßnahmen eingeleitet, um dieses zu verhindern. Der RTI-Ansatz ist damit ein *Präventionsmodell* (Gresham et al., 2005; Hale, Kaufman, Naglieri & Kavale, 2006; Hartmann, 2008; NASDSE, n. d.; Vaughn & L. S. Fuchs, 2003).

Ein weiteres Ziel des RTI-Ansatzes ist die bessere Kooperation zwischen allgemeiner Pädagogik und Sonderpädagogik. Ziel ist es, Interventionen an die Art, den Umfang und die Intensität der Lernschwierigkeiten anzupassen. Dabei werden pädagogische

Maßnahmen der allgemeinen Pädagogik und Sonderpädagogik integriert (Gresham et al., 2005; Vaughn & L. S. Fuchs, 2003).

Die Ziele und zugleich die Vorzüge des RTI-Ansatzes lassen sich nach Hartmann (2008) in Anlehnung an Speece, Case und Molloy (2003) sowie Vaughn und L. S. Fuchs (2003) wie folgt zusammenfassen: *„Unterrichtsoptimierung; Vermeidung von ,teaching disabilities‘; Früherkennung und Prävention von Lernstörungen; Reduktion von diagnostischen Fehlklassifikationen; Verminderung des Anteils an Schulkindern in sonderpädagogischen Programmen; Überwindung der Kluft zwischen Diagnostik und Intervention; Verbesserung der Kooperation von Regel- und Sonderpädagogik“* (Hartmann, 2008, S. 127).

Der RTI-Ansatz beruht auf der Grundannahme, alle Schüler effektiv unterrichten zu können. Das bedeutet, dass Lehrkräfte die Verantwortung dafür tragen, eine Lernumgebung zu schaffen, in der alle Schüler adäquat lernen können. Für die Erreichung dieses Ziels ist die Früherkennung von Lern- und Verhaltensschwierigkeiten notwendig. Erst dann können präventive Interventionen eingeleitet werden, um Lernstörungen zu verhindern. Welche grundlegenden Modelle es zur Umsetzung von Interventionen im Rahmen des RTI-Ansatzes gibt, zeigt der nachfolgende Abschnitt auf.

2.4.4 Aufbau und Interventionsebenen des RTI-Modells

RTI-Modelle treffen keine Aussagen über zugrunde liegende Ursachen von Lernschwierigkeiten. RTI-Modelle verdeutlichen, dass Lernschwierigkeiten im Kind selbst und/oder im Unterricht liegen können. Im RTI-Ansatz wird deshalb durch eine systematische Stärkung der Unterrichtsqualität überprüft, ob Beeinträchtigungen des Kindes zu Lernschwierigkeiten führen (Speece et al., 2003). Es gibt nicht *das* RTI-Modell, sondern es gibt verschiedene Varianten von RTI-Modellen (Hartmann, 2008). Jedes RTI-Modell ist ein *Mehrebenenmodell*, das aus zwei bis vier Interventionsebenen besteht, wobei das 3-Ebenenmodell das bevorzugte Modell ist (L. S. Fuchs & D. Fuchs, 2007; Reschly, 2005; Vaughn, Wanzek, Woodruff & Linan-Thompson, 2007). Die Interventionsebenen sind hierarchisch aufeinander aufgebaut und unterscheiden sich in der Intensität der Intervention und in der Intensität der Lernmessungen (Reschly, 2005).

Innerhalb jeder Interventionsebene sind die Kooperation mit den Eltern und die Lernverlaufsdokumentation von großer Bedeutung. Nachfolgend wird der Aufbau des 3-Ebenenpräventionsmodells erläutert.

Die erste Interventionsebene entspricht der primären Prävention und beinhaltet den regulären Unterricht. Sie umfasst alle Schüler. Innerhalb der ersten Ebene wird die Lernentwicklung von Schülern durch Screenings, Klassenarbeiten oder durch systematische Lernverlaufsdokumentationen erhoben. Schüler mit leichten Lernschwierigkeiten erhalten zusätzliche Förderung in kleinen Gruppen innerhalb des regulären Unterrichts. Schüler, für deren Lernentwicklung der reguläre Unterricht nicht ausreichend ist und die anfällig für Lernschwierigkeiten sind, kommen in die zweite Interventionsebene (L. S. Fuchs & D. Fuchs, 2007; Hartmann, 2008; Hartmann & C. Müller, 2009; National Center for Learning Disabilities, 2006; National Joint Committee on Learning Disabilities, 2005; Reschly, 2005).

Die zweite Interventionsebene entspricht der sekundären Prävention und umfasst alle Risikoschüler. Innerhalb der zweiten Ebene wird die Lernentwicklung von Schülern durch Lernverlaufsdokumentationen erhoben, um die Effektivität der Interventionen zu beurteilen und Interventionen und schulische Ressourcen den Bedürfnissen des Schülers anzupassen. Die zweite Interventionsebene beinhaltet zum regulären Unterricht ergänzende, fokussierte Interventionen sowie die Kooperation mit weiteren Lehrkräften und/oder Spezialisten. Fokussierte Interventionen werden zusätzlich zum regulären Unterricht eingesetzt und sind auf curriculare Lernziele gerichtet. Sie finden mehrmals täglich oder wöchentlich in Kleingruppen oder als Einzelinterventionen innerhalb oder außerhalb des regulären Unterrichts statt und erstrecken sich über 15 bis 20 Wochen. Schüler, für deren Lernentwicklung fokussierte Interventionen ausreichen und die die Lernziele der Klasse wieder erreicht haben, nehmen nicht mehr an der zusätzlichen Förderung teil. Schüler, für deren Lernentwicklung fokussierte Interventionen nicht ausreichen, kommen in die dritte Interventionsebene (L. S. Fuchs & D. Fuchs, 2007; Hartmann, 2008; Hartmann & C. Müller, 2009; National Center for Learning Disabilities, 2006; National Joint Committee on Learning Disabilities, 2005; Reschly, 2005). Es gibt zwei Möglichkeiten, fokussierte Interventionen durchzuführen: nach dem *Problem solving-model* (Problemlösungsmodell) und dem *Standard protocol-model* (Standard-Protokollmodell). Beide Modelle unterscheiden sich in der Individualisierung der

Intervention. Das Problemlösungsmodell beinhaltet individualisierte und flexible Interventionen. Es ist ein bevorzugtes Vorgehen von Lehrkräften, um die Lernleistungen schwacher Schüler individuell zu steigern. Das Standard-Protokollmodell beinhaltet standardisierte und evidenzbasierte Interventionen für Schülergruppen mit ähnlichen Lernschwierigkeiten. Es ist ein strukturiertes Vorgehen mit remedialen Charakter von ca. 10 bis 15 Wochen, konzipiert für kleinere Gruppen von 6 bis 10 Schülern (D. Fuchs & L. S. Fuchs, 2006; D. Fuchs, Mock, Morgan & Young, 2003; Hartmann; National Research Center on Learning Disabilities, 2007; Shapiro, 2009). Während das Problemlösungsmodell sensitiv für Lernschwierigkeiten und -bedürfnisse ist, liegt der primäre Vorteil des Standard-Protokollmodells in dem standardisierten Vorgehen, welches eine bessere Kontrolle der Wirksamkeit der Intervention erlaubt. Ein weiterer Vorteil ist, dass Schulen eine kleine Anzahl effektiver Interventionsmaßnahmen für mehrere Schüler mit ähnlichen Lernschwierigkeiten verwenden können. Dadurch ist eine effiziente Nutzung von schulischen Ressourcen möglich. Zugleich stellt das Standard-Protokollmodell aber eine Herausforderung für Interventionen für Schüler mit komplexen Lernschwierigkeiten dar, da diese Schüler nicht allgemein angelegte Interventionen, sondern individualisierte Interventionen benötigen (D. Fuchs et al., 2003; Hartmann, 2008; Shapiro, 2009).

Die dritte Interventionsebene entspricht der tertiären Prävention und beinhaltet intensive, systematische, spezielle, sonderpädagogische Maßnahmen sowie eine vertiefte Diagnostik. Dieses Vorgehen entspricht zum Teil dem Vorgehen zur Feststellung des sonderpädagogischen Förderbedarfs. Innerhalb der dritten Ebene wird die Lernentwicklung von Schülern durch mehrmals wöchentlich stattfindende Lernverlaufsdokumentationen erhoben. Die dritte Interventionsebene wird außerhalb des regulären Unterrichts von Sonderpädagogen und anderen Fachleuten für ein bis zwei Schüler geleitet und erstreckt sich über einen sehr langen Zeitraum von bis zu mehreren Jahren. Die Schüler werden in den regulären Unterricht entlassen, wenn die speziellen Interventionen für ihre Lernentwicklung ausreichen und sie die Lernziele ihrer Klassen erreichen. Sollten dann erneut Lernschwierigkeiten auftreten, werden erneut fokussierte/spezielle Interventionen angeboten (L. S. Fuchs & D. Fuchs; Hartmann, 2008; Hartmann & C. Müller, 2009; National Center for Learning Disabilities, 2006; National Joint Committee on Learning Disabilities, 2005; Reschly, 2005). Abbildung 2 fasst das 3-Ebenenpräventionsmodell zusammen.

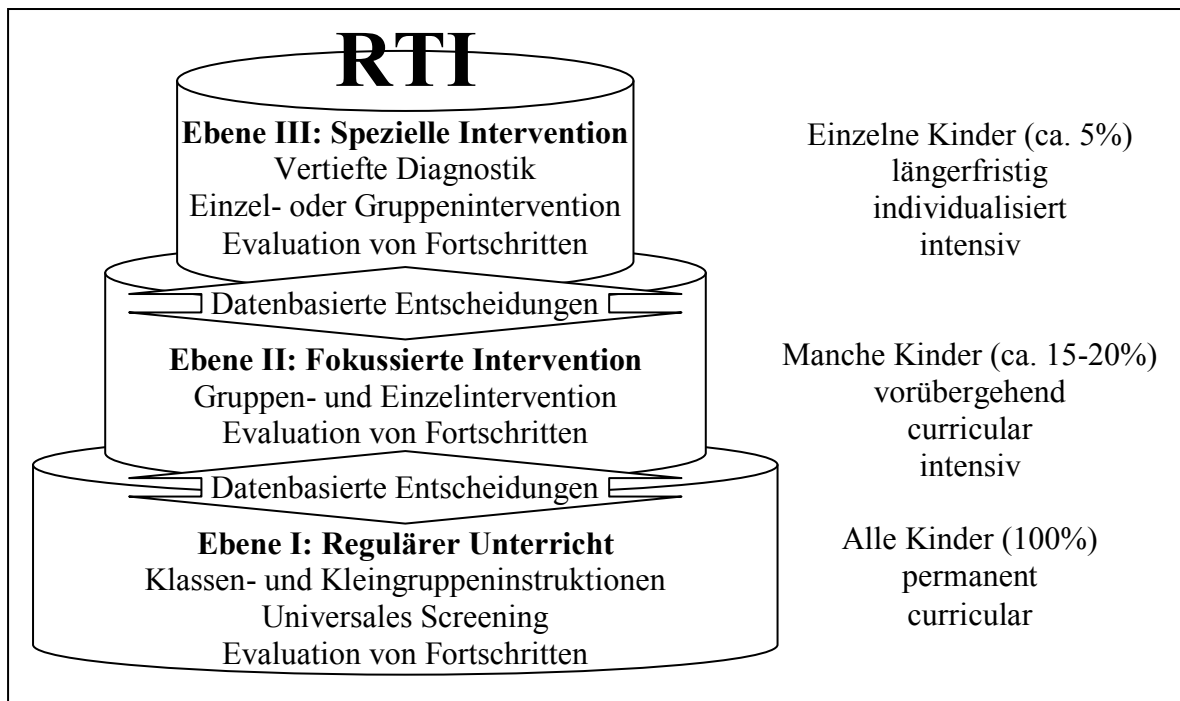


Abbildung 2: RTI-Modell (Hartmann & C. Müller, 2009, S. 27)

Alle unterschiedlichen RTI-Modelle sind durch Kernelemente miteinander verbunden. Zu den Kernelementen zählen:

1. evidenzbasierte Interventionen und der damit verbundene Anspruch, alle Kinder nach qualitativ hochwertigen Maßstäben zu unterrichten,
2. valide Messungen von antwortbezogenem Verhalten auf den Unterricht und der damit verbundene Anspruch, durch regelmäßige Messungen Unterricht zu optimieren,
3. datenbasierte Anpassungen des Unterrichts und der Interventionen und der damit verbundene Anspruch, den individuellen Lernbedürfnissen der Schüler gerecht zu werden (National Joint Committee on Learning Disabilities, 2005).

Im folgenden Abschnitt wird als Grundlage zur Entwicklung der FE-RS 2 ein wesentliches Kernelement aller RTI-Modelle, die Messverfahren zur Lernverlaufsdokumentation, erläutert.

2.4.5 Verfahren zur Lernverlaufsdokumentation: Mastery Measurement und General Outcome Measurement

Der wiederholte Einsatz von Messverfahren zur Erfassung antwortbezogenen Verhaltens von Schülern auf Unterricht und Interventionen (RTI) liefert Daten zum allgemeinen Lernverlauf oder zur Entwicklung spezifischer Fähigkeiten von Schülern. Die

Arbeitsergebnisse der Schüler und damit erstellbare Verlaufsprotokolle des Antwortverhaltens können analysiert werden. Die wiederholte Anwendung der Messverfahren hat auf der Grundlage der erhobenen Daten das Ziel, die Effektivität von Interventionen zu beurteilen sowie nachfolgend Unterricht und Intervention den Bedürfnissen des Schülers anzupassen. Die Gestaltung von Unterricht und Intervention im RTI-Ansatz ist damit ein datenbasierter Prozess. Messverfahren dieser Art sind unter dem Begriff *Progress Monitoring* zusammengefasst (Salvia & Ysseldyke, 2007). In der vorliegenden Arbeit werden sie als Verfahren zur Lernverlaufsdokumentation übersetzt.

Verfahren zur Lernverlaufsdokumentation sind effektive, zeitökonomische und formative Messverfahren für die allgemeine Pädagogik und Sonderpädagogik. Die Verfahren dienen der Erfassung von Leistungsveränderungen über einen bestimmten Zeitraum. Sie ermöglichen, Schüler zu identifizieren, die keine angemessenen Lernfortschritte zeigen und deshalb zusätzliche oder alternative Instruktionen/Interventionen benötigen. Daneben erlauben Verfahren zur Lernverlaufsdokumentation auch Formulierungen von Aussagen über die Wirksamkeit von Unterricht und Interventionen, das heißt darüber, ob der Schüler vom Unterricht profitiert und damit ein *Responder* ist oder ob veränderte Instruktionen notwendig sind, da der Schüler ein *Non-Responder* ist. Durch den datenbasierten Vergleich der Wirksamkeit unterschiedlicher Instruktionen/Interventionen ermöglichen sie den Lehrkräften, Unterricht effizient und binnendifferenziert zu gestalten. Kennzeichnend für diese Messverfahren sind regelmäßige, sich wiederholende kurze Lernverlaufsdokumentationen, die mehrmals in der Woche stattfinden und eine bis fünf Minuten dauern können. Bei der Lernverlaufsdokumentation werden zunächst der aktuelle Leistungsstand des Schülers in einem bestimmten Bereich erhoben und das Leistungsniveau (Ziellinie/Aimline) festgelegt. Anschließend wird in regelmäßigen Abständen mit ähnlichen Aufgabenstichproben die Leistung des Schülers gemessen. Die Lernentwicklung wird erfasst, indem der jeweils aktuelle Leistungsstand mit dem vorherigen Leistungsstand und mit dem erwarteten Leistungsniveau verglichen wird. Um die Lernentwicklung des Schülers oder der Klasse übersichtlich darzustellen, werden die erhobenen Daten in Diagramme überführt (L. S. Fuchs & D. Fuchs, n. d.; Salvia & Ysseldyke, 2007). Abbildung 3 zeigt ein allgemeines Diagramm zur Leistungsentwicklung.

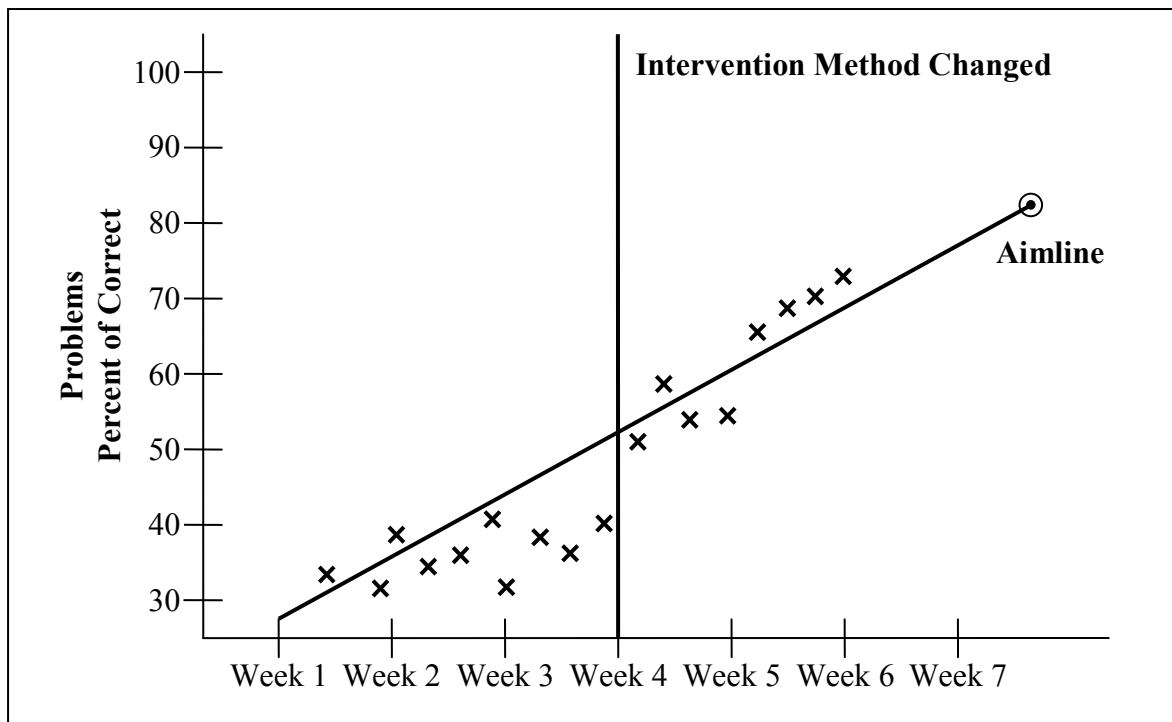


Abbildung 3: A Progress-Monitoring Chart with an Aimline (Salvia & Ysseldyke, 2007, S. 648)

In Abbildung 3 wird ein Lernverlauf über 7 Wochen dargestellt. Die Ordinate gibt die Wochenanzahl und Häufigkeit der Messwiederholungen, die Abszisse die prozentuale Häufigkeit der richtigen Antworten. Die Ziellinie kennzeichnet das zu erreichende Leistungsniveau nach 7 Wochen. In den ersten 4 Wochen liegt der erfasste Leistungsstand überwiegend unterhalb der Ziellinie. Nach einem Wechsel der Intervention in der 4. Woche steigert sich das Leistungsniveau und der erfasste Leistungsstand liegt überwiegend oberhalb der Ziellinie.

Es gibt zwei Möglichkeiten festzustellen, ob die Intervention den Lernbedürfnissen des Schülers entspricht und damit die Leistungsentwicklung des Schülers adäquat verläuft: die *4-Punkte-Regel* und die *Trendlinien-Regel*. Liegen nach der 4-Punkte-Regel die letzten vier Messergebnisse über der Ziellinie, gelten die Intervention und die Leistungsentwicklung als adäquat. Liegen alle vier Messergebnisse unter der Ziellinie, sollte eine Anpassung der Intervention erfolgen. Nach der Trendlinien-Regel wird der Steigungsfaktor der letzten acht Messergebnisse, die die Trendlinie bilden, berechnet. Ist die Steigung der Trendlinie steiler als die der Ziellinie, gelten die Intervention und die Leistungsentwicklung als adäquat. Ist die Steigung der Trendlinie flacher als die der Ziellinie, sollte eine Anpassung der Intervention erfolgen (McLane, n. d.).

Es gibt zwei Arten von Verfahren der Lernverlaufsdokumentation: *Mastery Measurement* (MM) und *General Outcome Measurement* (GOM) (L. S. Fuchs & Deno, 1991).

MM ist ein aufgabenanalytisches Vorgehen. Beispiele für MMs sind *Curriculum-Based Evaluation* (CBE) und *Curriculum-Based Assessment* (CBA). Beim MM-Vorgehen wird eine komplexe Fähigkeit in mehrere Teilfähigkeiten unterteilt. Die Teilfähigkeiten werden sodann als *kurzfristig* erreichbare Lernziele hierarchisch geordnet. Auf Grundlage der hierarchischen Aufgabenanalyse werden MM-Verfahren und Items erstellt, die die Teilfähigkeiten der Schüler überprüfen können. Für den Lernbereich Lesen würde z. B. die komplexe Fähigkeit des Lesens in Phonemerkennung, Rekodieren, Dekodieren, etc. differenziert werden. Die Teilfähigkeiten würden sodann geordnet und entsprechend dieser Reihenfolge im Unterricht vermittelt werden. Entsprechend der gleichen Hierarchie würde ein MM-Verfahren erstellt werden und in regelmäßigen Zeitabständen die Teilfähigkeiten überprüfen. Erst wenn durch den Einsatz des MM-Verfahrens deutlich wird, dass die Schüler die ersten, die hierarchisch untersten, Teilfähigkeiten beherrschen, würden die nächsten Lernziele im Unterricht angestrebt werden. Vorteil des MM-Vorgehens ist, dass die Lehrplan- und Unterrichtsvalidität gesichert ist, denn Lehrkräfte überprüfen nur die Fähigkeiten, die im Lehrplan verankert sind und im Unterricht vermittelt werden. Gleichzeitig ist dies ein Nachteil des MM-Vorgehens. MM-Verfahren berücksichtigen nur aktuelle Lerninhalte, aber nicht zurückliegende Lerninhalte, was zu Problemen bei Schülern mit Lernschwierigkeiten führen kann. Ebenso führt dieses nach kurzfristigen Lernzielen ausgerichtete Vorgehen zu Problemen bei der Dokumentation der Lernentwicklung, denn sobald ein Schüler Lernfortschritte in einer Teilfähigkeit zeigt und das Leistungsniveau steigt, wird eine neue Teilfähigkeit vermittelt, wodurch das Leistungsniveau dieser neuen Teilfähigkeit auf ein unteres Level sinkt. Die in einem Diagramm dargestellten Daten würden dann wie Sägezähne aussehen; eine allgemeine Lernentwicklung wäre nicht ablesbar. Ein wichtiger Aspekt ist, dass das MM-Vorgehen auf nicht bestätigten Annahmen über die Reliabilität und Veränderungssensitivität des Verfahrens beruht. Ein weiterer Nachteil ist, dass die Praktikabilität des Vorgehens durch die hohe Anzahl benötigter MM-Verfahren, die der Anzahl der Teilfähigkeiten entspricht, eingeschränkt ist. Lehrkräfte können dadurch leicht den Überblick über die Vielzahl unterschiedlicher MM-Verfahren verlieren (L. S. Fuchs & Deno, 1991).

GOM ist ein valides, reliables und effizientes Vorgehen für datenbasierte

Lernverlaufsdokumentationen und Unterrichtsentscheidungen. Beispiele für GOMs sind *Curriculum-Based Measurement* (CBM) und *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS). Das GOM-Vorgehen fokussiert komplexe Fähigkeiten und *langfristige* Lernziele. Die Lernverlaufsdokumentation erfolgt mit wiederholt stattfindenden Aufgabenstichproben und Items, die einer komplexen Fähigkeit und einem langfristigen Lernziel entsprechen. Für den Lernbereich Lesen würde z. B. die komplexe Fähigkeit des Lesens mit dem lauten Lesen eines Textes überprüft werden. Vorteil des GOM-Vorgehens ist, dass komplexe und umfassende Lernziele überprüft werden. Zum einen wird damit die schwierige und zeitintensive Aufgabe vermieden, komplexe Lernziele zu zergliedern und zum anderen ermöglicht es, die Wirksamkeit des Unterrichts besser zu beurteilen. Im Vergleich zum MM-Vorgehen, das die Reihenfolge der Lerninhalte vorschreibt, ermöglicht das GOM-Vorgehen den Lehrkräften, unterschiedliche Unterrichtsmethoden, Verfahren und Materialien anzuwenden und aufgrund der erhobenen Daten die Wirksamkeit und Effektivität dieser zu beurteilen. Unterricht (unabhängige Variable) und Leistungsmessungen (abhängige Variable) sind also nicht wie beim MM-Vorgehen miteinander verwoben. Nachteil des GOM-Vorgehens ist die geringere Veränderungssensibilität aufgrund der Komplexität und Langfristigkeit der Lernziele. Trotz dieser Einschränkung erlauben GOM-Verfahren sensitive und datenbasierte Entscheidungsfindungen (L. S. Fuchs & Deno, 1991). Nachfolgend werden die Verfahren der Lernverlaufsdokumentation CBE, CBA, CBM und DIBELS kurz beschrieben und miteinander verglichen.

Sind Verfahren *Curriculum-Based*, stehen sie in direktem Zusammenhang zu dem, was im Unterricht an Wissen, Fertigkeiten und Fähigkeiten vermittelt wird. Sie basieren auf den gültigen Lehrplaninhalten und den im Unterricht verwendeten Materialien.

Ein CBE-Verfahren ist „a thoughtful process of comparison and judgment“ (Howell, Kurns & Antil, 2002, S. 756). Es ist ein strukturiertes Vorgehen, um unterrichtliche Entscheidung vor dem Hintergrund der Schülerleistungen zu treffen und um festzustellen, ob der Unterricht den gewünschten Erfolg bringt. Zudem ist das Verfahren ein diagnostisches Instrument, um hypothesengeleitet Aussagen über Lernschwierigkeiten treffen zu können. Bei CBE-Verfahren liegt der Fokus jedoch nicht auf verschiedenen Messinstrumenten, sondern auf Wenn-Dann-Aussagen, die Grundlagen aller unterrichtlichen Entscheidungsprozesse sind. Deshalb haben CBEs keine eigenständigen Messinstrumente,

um Daten über Unterricht und Leistungen zu erheben, sondern nutzen CBMs, die darauf ausgelegt sind, zu messen. Die Grundannahme des CBE-Verfahrens in allen Entscheidungsprozessen ist, dass wenn ein Schüler etwas nicht tun kann, dann deshalb, weil er nicht weiß wie. Das heißt, dem Schüler fehlen wesentliche Fähigkeiten, um eine bestimmte Aufgabe zu lösen. Die Wenn-Dann-Aussagen formulieren Hypothesen über die Leistungen eines Schülers und werden im CBE-Prozess sodann in ein Flussdiagramm überführt, welches versucht, durch Fragen, Wenn-Dann-Aussagen und diagnostische Verfahren potenziell fehlende Fähigkeiten einzugrenzen und mögliche Interventionsvorschläge zu unterbreiten. Stellt sich beispielsweise eine Lehrkraft die Frage, ob die Lesefähigkeit eines Schülers zufriedenstellend ist, nutzt die Lehrkraft z. B. Interviews, Klassenarbeiten, zugängliche Vorinformationen oder CBMs. Wenn die Antwort aufgrund der diagnostischen Ergebnisse *ja* lautet, dann wird dieser CBE-Prozess beendet und es kann ein neuer CBE-Prozess für eine andere Fähigkeit, z. B. das Leseverständnis, begonnen werden. Wenn die Antwort aufgrund der diagnostischen Ergebnisse *nein* lautet, dann wird eine nächste, spezifischere Frage gestellt, z. B. ob der Schüler einige wenige Wörter lesen kann. Diese Frage wird anschließend solange beantwortet oder spezifiziert, bis eine Intervention für den Schüler vorgeschlagen werden kann, z. B. die phonologische Bewusstheit zu trainieren (Howell et al., 2002).

CBA-Verfahren sind jegliche informelle curriculumbasierte Testverfahren (Deno, 2003). Es gibt zwei Arten von Modellen von CBAs: *accuracy-based models* (genauigkeitsbasiertes Modell, z. B. Gickling & Havertape, 1981) und *criterion-referenced models* (kriteriumsorientiertes Modell, z. B. Blankenship, 1985). Das CBA-Modell von Gickling und Havertape (1981) soll Lehrkräfte bei Unterrichtsplanungen unterstützen und Unterschiede vermeiden, die zur Nichtpassung zwischen Schüler und Lernumgebung führen. Gickling und Havertape (1981) formulieren das Hauptanliegen ihres Modells wie folgt: „The curriculum based assessment model provides a structure for controlling the differences between what the teacher is trying to teach and what the student is able to learn“ (S. 220). Ziel ist es, aufgrund datenbasierter Entscheidungen die Lernleistung der Schüler zu steigern. Die wichtigste Grundlage der datenbasierten Entscheidungen ist dabei der genaue Prozentsatz richtiger Antworten eines Schülers. Deshalb ist dieses Modell ein genauigkeitsbasiertes Modell. Folgendes Beispiel erläutert dieses Verfahren: Ein Schüler soll 1 Minute lang einen Text laut lesen. Anschließend werden die Anzahl und der Prozentsatz richtig gelesener Wörter berechnet. Je nach erreichtem Prozentsatz hat der

Schüler dann im 1-Minute-Lesen den Leistungsstand *frustration level* (< 93 %), *instructional level* (93 %-97 %) oder *independent level* (> 97 %) erreicht. Entsprechend des erreichten Prozentsatzes können Unterrichtsprozesse angepasst werden (Gickling & Havertape, 1981). Das CBA-Modell von Blankenship (1985) soll ähnlich wie das Modell von Gickling und Havertape (1981) Lehrkräfte bei Unterrichtsplanungen unterstützen. „The essence of the approach is the linking of assessment to curriculum and instruction“ (Blankenship, 1985, S. 234). Ziel dieses Modells ist es, durch die Dokumentation und Einschätzung unterschiedlicher Fähigkeiten geeignete Lehrinhalte zu bestimmen. Um geeignete Lehrinhalte, dazu gehören auch Materialien wie Schulbücher, für jeden Schüler auszuwählen, werden CBAs in der Regel am Schuljahresanfang eingesetzt. CBAs können sich auf ganz unterschiedliche Inhalte des Curriculums und individuelle Lernziele beziehen. Werden ähnliche Leistungen wiederholt und regelmäßig überprüft, können daraus konkrete Hinweise über die Leistungen der Schüler, über ihr Erreichen von Lernzielen und über anschließend geeignete Lehrziele abgeleitet werden. Damit dient CBA nicht der Erfassung individueller Differenzen, sondern der Leistungseinschätzung eines Schülers bezüglich eines spezifischen Lehrziels. Deshalb ist dieses Modell ein kriteriumsorientiertes Modell. Folgendes Beispiel erläutert dieses Verfahren: Eine Lehrkraft möchte das Thema „Längen“ im Mathematikunterricht behandeln und diesbezüglich mittels eines CBAs geeignete Lehrinhalte für jeden Schüler auswählen. Für die Entwicklung dieses CBAs nutzt sie das Mathematiklehrwerk. Sie listet alle im Lehrwerk geforderten und themenbetreffenden Fähigkeiten auf und ordnet diesen Aufgaben zu. Entsprechend der gelösten Aufgaben bestimmt die Lehrkraft für jeden Schüler das aktuelle Leistungsniveau und die Lehrziele und plant die Unterrichtseinheit. Nach einigen Unterrichtsstunden zum Thema lösen die Schüler erneut ein CBA mit ähnlichen Aufgaben. Entsprechend der gelösten Aufgaben erhält die Lehrkraft Hinweise auf den Lernverlauf der Schüler, bestimmt erneut die Lehrziele und passt den Unterricht an (Blankenship, 1985).

CBM ist „a process of measurement and data collection“ (Howell et al., 2002, S. 756). Es ist ein Set standardisierter, valider, kurzer und sensibler Tests, welches, in regelmäßigen Zeitabständen wiederholt eingesetzt, die Lernentwicklung von Schülern dokumentiert und Rückmeldung über die Wirksamkeit des Unterrichts gibt. CBMs werden vorwiegend für Basisfertigkeiten wie Lesen, Schreiben, Rechtschreiben und Rechnen verwendet (Deno, 2003). „CBM-Verfahren erlauben also vor allem Aussagen zum Fortschritt der

Automatisierung und nicht über zugrunde liegende basale Verständnisvoraussetzungen. [Es] wird eine ergebnisorientierte Fortschrittserhebung angestrebt, die Hinweise auf Schwierigkeiten und Bedürfnisse einzelner Kinder gibt, welche individuell aufgearbeitet werden müssen“ (C. Müller & Hartmann, 2009, S. 39). Ziel des Verfahrens ist es, den Unterricht den Bedürfnissen der Schüler anzupassen und dadurch den Unterricht zu optimieren. CBM-Tests können von Lehrkräften auf Grundlage der Lehrpläne und der im Unterricht verwendeten Materialien konzipiert werden. Für die selbstständige CBM-Aufgabenentwicklung gibt es Leitfäden (z. B. „The ABCs of CBM“ von M. K. Hosp, J. L. Hosp und Howell [2007]). Des Weiteren können im Internet mit Hilfe eines Zufallsgenerators konzipierte CBMs (z. B. unter www.interventioncentral.org) oder bereits bestehende CBMs (z. B. unter www.easycbm.com) heruntergeladen werden. Außerdem können CBMs auch käuflich erworben werden (z. B. unter www.aimsweb.com). Das folgende Beispiel erklärt ein CBM: Bei einem Rechtschreib-CBM werden 2 Minuten lang einzelne Wörter diktiert, die die Schüler schreiben sollen. Homophone werden für ein besseres Sinnverständnis im Satz diktiert. Für die erste und zweite Klasse werden 12 Wörter (alle 10 Sekunden ein neues Wort), ab der 3. Klasse 17 Wörter (alle 7 Sekunden ein neues Wort) diktiert. Die Anzahl orthographisch richtig geschriebener Wörter und die Anzahl richtiger Zwischenstellen zwischen den Graphemen eines Wortes (correct letter sequences [CLS]) (z. B. fünf CLS in $^h^e^e^l^$) sind hierbei Maße für die Rechtschreibleistung (M. R. Shinn & M. M. Shinn, 2002). Abbildung 4 zeigt ein Beispiel für ein Rechtschreib-CBM der 3. Klassenstufe.

AIMSweb® Standard Spelling Benchmark Assessment List #1 (3rd Grade)			
Given By: _____		Date Given: ____/____/____	
ID	Word	CLS	CCLS
1	heel The heel of the shoe.	5	5
2	wise	5	10
3	face	5	15
4	speak	6	21
5	dinner	7	28
6	weren't Weren't there better things to do?	8	36
7	there It is over there!	6	42
8	river	6	48
9	nameless	9	57
10	able	5	62
11	tomorrow	9	71
12	highway	8	79
13	mess	5	84
14	headaches	10	94
15	wag	4	98
16	crown	6	104
17	scraped	8	112
Total CLS			112

Copyright 2003 Edformation, Inc.
www.AIMSweb.com

Abbildung 4: Beispiel für ein Rechtschreib-CBM der 3. Klassenstufe (AIMSweb, 2003)

Anmerkungen: CLS = correct letter sequences, CCLS = cumulative number of correct letter sequences

DIBELS sind Sets von kurzen standardisierten Messverfahren, um den frühen Erwerb grundlegender schriftsprachlicher Fähigkeiten vom Kindergarten bis zur 6. Klassenstufe zu beurteilen. Sie beruhen auf den Ideen der CBM-Verfahren, stützen sich aber nicht ausschließlich auf Lehrplanvorgaben. DIBELS werden dreimal im Schuljahr (im Herbst, Winter und Frühling) durchgeführt und beanspruchen pro Schüler 5 bis 10 Minuten. Ziel von DIBELS ist es, Schüler mit Schwierigkeiten im frühen Schriftspracherwerb zu erkennen und zu fördern, die Effektivität der Interventionen einzuschätzen und gegebenenfalls zu ändern sowie die Lernentwicklung der Schüler zu dokumentieren und zu stärken. DIBELS erheben nicht Daten zu allen schriftsprachlichen Fähigkeiten, sondern messen nur *Indikatoren* für frühe schriftsprachliche Fähigkeiten. So werden z. B. nicht alle Facetten der phonologischen Bewusstheit, sondern nur die Anlautidentifikation und Phonemsegmentation als Indikatoren für phonologische Bewusstheit erfasst. Insgesamt werden folgende sieben Indikatoren für schriftsprachliche Fähigkeiten erhoben: Anlautidentifikation, Buchstaben benennen, Phonemsegmentation, Pseudowörter lesen, lautes Lesen, Leseverständnis und Satz mit vorgegebenem Wort bilden (Kaminski,

Cummings, Powell-Smith & Good III, 2008).

Curriculumbasierte Messverfahren dokumentierten Lernverläufe von Schülern durch Unterricht und Interventionen (RTI). Sie sind Verfahren, die Schüler mit Lernschwierigkeiten frühzeitig identifizieren und die die Wirksamkeit von Unterricht und Intervention überprüfen. Es gibt zwei Arten von Verfahren zur Lernverlaufsdokumentation: Zum einen MM-Verfahren, zu denen CBE und CBA zählen, die kurzfristige Lernziele erfassen und zum anderen GOM-Verfahren, zu denen CBM und DIBELS zählen, die langfristige Lernziele erfassen. Einschränkend ist hinzuzufügen, dass allein die Anwendung von Verfahren zur Lernverlaufsdokumentation nicht zur Optimierung des Unterrichts beiträgt (Capizzi & L. S. Fuchs, 2005). „Durch wiederholte Messungen allein entstehen keine Lernfortschritte, ebenso wenig durch eine Modifikation von Förderentscheidungen über eine ‚Versuchs- und Irrtumsstrategie‘. . . . So gesehen fungiert CBM nur als ein Impulsgeber . . .“. (Diehl & Hartke, 2007, S. 202).

2.4.6 Kritik des RTI-Ansatzes

Der RTI-Ansatz ist ein empirisch abgesichertes Modell formativer Evaluation im Unterricht, mit dessen Hilfe ein „Wait-to-fail“-Vorgehen überwunden werden kann. Es betont die Prävention und frühe Diagnostik von Lernschwierigkeiten durch kontinuierliche und kleinschrittige Erfassung von Leistungsveränderungen von Schülern. Schüler, die Risiken unterliegen, die Lernschwierigkeiten bedingen können, werden frühzeitig identifiziert. Sie erhalten damit so früh wie möglich auf ihre individuellen Lernbedürfnisse abgestimmte Interventionen. Dieser RTI-Prozess beginnt bereits in Kindertagesstätten und begünstigt die Vermeidung von Lernschwierigkeiten. Die Wirksamkeit der Interventionen wird durch curriculumbasierte Messverfahren erhoben, wobei das Ziel der systematischen Stärkung und Optimierung des Unterrichts verfolgt wird (Hartmann & C. Müller, 2009; Vaughn & L. S. Fuchs, 2003; Walter, 2008a). Trotz der bemerkenswerten Ansätze des RTI-Ansatzes gibt es offene Fragen, vor allem hinsichtlich der Umsetzung in die Praxis, die im Folgenden skizziert werden.

Gibt es im RTI-Ansatz eine „echte“ LD? Traditionell wird Behinderung als ein Defizit innerhalb des Individuums definiert, das durch Kontextvariablen beeinflusst, aber nicht

verursacht wird. Der RTI-Ansatz beschreibt keine kognitiven Defizite. Zweck des RTI-Ansatzes ist es, Kontextvariablen als mögliche Ursachen für Lernschwierigkeiten auszuschließen. Gelingt dieser Ausschluss von Kontextvariablen nicht, das heißt, kann ein Schüler nicht auf valide und evidenzbasierte, intensive Interventionen adäquat reagieren, liegt das Defizit innerhalb des Individuums. Es liegt damit eine „echte“ LD vor (Vaughn & L. S. Fuchs, 2003).

Gibt es valide Interventionsmodelle und Messverfahren, um die Wirksamkeit von Unterricht und Interventionen sicherzustellen? Am weitesten ist der Forschungsstand zu validen und reliablen Interventionsmodellen und Messverfahren im Bereich Lesen voran geschritten. Hier konnten in Studien hohe Validitäten von über $r = .60$ und mittlere und hohe Reliabilitäten zwischen $r = .80$ und $r = .94$ nachgewiesen werden. Für den Mathematikunterricht, die Bereiche Schriftlicher Ausdruck und Rechtschreiben gibt es ebenso valide und reliable Messverfahren, um den Lernverlauf zu dokumentieren. Insgesamt ist der Forschungsstand für den Vorschul- und Primarbereich weit voran geschritten. Für den Sekundarbereich und weitere Unterrichtsfächer bedarf es noch intensiver Forschungsanstrengungen (Hosp et al., 2007; Vaughn & L. S. Fuchs, 2003).

Was geschieht mit Schülern, die unzureichend auf spezielle, sonderpädagogische Maßnahmen reagieren? Reagieren Schüler inadäquat auf spezielle Interventionen, liegt das Defizit innerhalb des Individuums und nicht innerhalb des Unterrichts. Möglicherweise ist anhaltende Nichtresponsivität durch sekundäre Defizite wie Aufmerksamkeitsdefizite, sensorische, emotionale und motivationale Schwierigkeiten oder durch eine geistige Behinderung bedingt. Differenzialdiagnostische Untersuchungen sind notwendig, da eine reine RTI-Diagnostik nur eine heterogene Gruppe von Schülern identifiziert, die aus verschiedenen Gründen inadäquat auf spezielle Interventionen reagiert (Hartmann, 2008; Vaughn & L. S. Fuchs, 2003). In dieser erweiterten Diagnostikphase werden sonderpädagogische Maßnahmen, für welche die Eltern und ein interdisziplinäres Team ihr Einverständnis gegeben haben, mit Hilfe von CBM evaluiert. Durch den Einsatz von CBMs sollen Maßnahmen gefunden werden, die sowohl das Leistungsniveau als auch den Lernfortschritt steigern. Nach acht Wochen beraten das Team und die Eltern über die CBM-Daten und das weitere Vorgehen. Treten Verbesserungen ein, werden die speziellen Interventionen fortgeführt. Treten keine Verbesserungen ein, wird die Diagnostikphase fortgeführt und weitere Maßnahmen hinzugezogen. Auf diese Weise können Schüler

identifiziert werden, für die andere Curricula als die des allgemeinen Unterrichts angemessen sind (Vaughn & L. S. Fuchs, 2003).

Nach welchem Interventionsmodell sollte im RTI-Prozess verfahren werden? In der zweiten Ebene des RTI-Prozesses werden zwei Interventionsmodelle voneinander unterschieden: das Problemlösungsmodell und das Standard-Protokollmodell. Ziel des Problemlösungsmodells ist es, den allgemeinen Unterricht auf die individuellen Lernbedürfnisse eines Schülers anzupassen. Reagiert der Schüler auf diese Adaption positiv, wird die Förderung auf der zweiten Interventionsebene nicht fortgesetzt. Beim Standard-Protokollmodell wird in Kleingruppen nach einem festgelegten Programm gearbeitet. Reagiert der Schüler auf dieses Programm positiv, wird die Förderung auf der zweiten Interventionsebene nicht fortgesetzt. Allerdings müssen bei diesem Vorgehen nach einiger Zeit viele Schüler wieder fokussierte Interventionen erhalten. Welches Vorgehen in der zweiten Ebene für welche Schüler bevorzugt werden sollte, ist derzeit Gegenstand von Forschungsanstrengungen (Vaughn & L. S. Fuchs, 2003).

Gibt es ausreichend geschultes Personal für die Durchführung des RTI-Ansatzes? Für die Durchführung des RTI-Ansatzes braucht es Lehrkräfte, die Kenntnisse über valide Unterrichtsverfahren und Interventionen haben und die die CBM-Technik beherrschen. Für viele Regel- und Sonderpädagogen sowie Schulpsychologen erfordert der RTI-Ansatz ein Umdenken und einen Paradigmenwechsel. Eine großflächige Anwendung des Ansatzes erfordert entsprechende Ausbildungsmaßnahmen in der Lehrerbildung sowie Fortbildungs- und Trainingsmaßnahmen für praktizierende Lehrkräfte (Vaughn & L. S. Fuchs, 2003). Inwieweit Lehrkräfte und andere Professionelle auf diese neue Aufgabe in den USA systematisch vorbereitet werden, ist nicht bekannt.

2.5 Zusammenfassung und Schlussfolgerungen

Wie in diesem Kapitel herausgearbeitet, wurde der RTI-Ansatz in Auseinandersetzung mit Problemen des ATI-Ansatzes entwickelt. Im Rahmen des ATI-Ansatzes wurde über mehrere Dekaden hinweg versucht, stabile Wechselwirkungen zwischen Aptitudes und Treatments zu ermitteln. Es war das Ziel, für ausgewählte Gruppierungen von Schülermerkmalen die passende Unterrichtsmethode zu finden und Unterricht adaptiv zu

gestalten. Obwohl die ATI-Forschung von Beginn an von Schwierigkeiten geprägt war, bestimmt sie noch heute die Förderpädagogik und -diagnostik in Deutschland. Eine Förderpädagogik und -diagnostik, die auf den Ideen des ATI-Ansatzes beruht, bietet viele mögliche individuelle Adaptionsmaßnahmen, sodass schulische Prävention, Früherkennung von Lernschwierigkeiten und Binnendifferenzierung möglich wären. Da aber der ATI-Ansatz von statischen und validen Zuordnungen zwischen Aptitudes und Treatments ausgeht, die im Unterricht nicht vorzufinden sind, führt dieser Ansatz zu Schwierigkeiten und damit zu Problemen bei Prävention, Differenzierung und Individualisierung. Das Ziel einer adäquaten Passung von pädagogischen Handlungen und Aptitudes wird nicht erreicht. Ein weiteres wesentliches Problem des ATI-Ansatzes ist die Definition von Lernschwierigkeiten nach der Diskrepanzdefinition, da die Diskrepanzdefinition keine präventiven und frühzeitigen Fördermaßnahmen erlaubt. „Oft muss jahrelang gewartet werden, bis die Schriftsprachleistung genügend weit hinter der Intelligenz zurück liegt, damit die Diagnose Dyslexie bzw. Legasthenie gestellt werden kann, die den Betroffenen späten Zugang zu systematischen Fördermaßnahmen eröffnet“ (Hartman, 2008, S. 124). An diesem Dilemma setzt der RTI-Ansatz an. Der RTI-Ansatz ist ein alternatives Diagnostik- und Präventionsmodell, das Schüler mit Lernschwierigkeiten frühzeitig in ihren Lernprozessen unterstützt. Es zielt durch eine systematische Stärkung der Unterrichtsqualität darauf ab, auszuschließen, dass Unterricht ursächlich für Lernschwierigkeiten ist. Der RTI-Ansatz wendet sich damit von einer defizitorientierten Sichtweise gegenüber dem Kind ab und sucht die Gründe für Lernschwierigkeiten von Schülern zuerst in Unterrichtsprozessen. Dies ist ein wesentlicher Unterschied zum ATI-Ansatz, weil dieser Lernschwierigkeiten häufig als Diskrepanz zur Intelligenz definiert. Auch der RTI-Ansatz definiert nach Hartmann (2008) Lernschwierigkeiten als Diskrepanz, aber eben als Diskrepanz zu effektiven pädagogischen Maßnahmen und den daraus abgeleiteten Lernerwartungen. Innerhalb des RTI-Ansatzes nehmen die curriculumbasierten Verfahren zur Lernverlaufsdokumentation eine zentrale Bedeutung ein. Diese Verfahren erfassen in regelmäßigen Zeitabständen die Lernentwicklung der Schüler, sie identifizieren Schüler, die nur geringe Lernfortschritte machen und erlauben dadurch auch, die Effektivität des Unterrichts einzuschätzen. Sie können aber, genauso wie die Förderdiagnostik, nicht voraussagen, welche Interventionen für die Schüler effektiv oder ineffektiv sein werden, sondern nur, welche es aufgrund der erhobenen Daten schon waren. Ziel dieses Vorgehens ist, Schüler mit Lernschwierigkeiten so früh wie möglich zu fördern und nicht im traditionellen Sinn umzuplatzieren bzw. zu selektieren. „Wenn eine

Selektion stattfindet, dann zwischen *Maßnahmen*, die sich als tauglich oder als untauglich erweisen“ (Walter, 2008a, S. 212).

Für die Entwicklung der FE-RS 2 ergibt sich aus den bisherigen Ausführungen Folgendes: Die formative Evaluation, ein wichtiges Merkmal des RTI-Ansatzes, wird lernprozess- und entwicklungsbegleitend durchgeführt und prüft den Erfolg der laufenden pädagogischen Maßnahmen, z. B. Unterricht oder Interventionen. Sie hat zum Ziel, durch objektive, valide, reliable und veränderungssensitive Lernverlaufsdokumentationen das inhaltliche und methodische Vorgehen pädagogischer Maßnahmen und deren Wirkung auf die Schüler zu bewerten. Die formative Evaluation gibt damit eine engmaschige Rückmeldung zur Effektivität und Wirksamkeit dieser Maßnahmen. Vorteil dieses Vorgehens ist es, dass durch die kontinuierliche und direkte Rückmeldung Lehrkräfte schon sehr frühzeitig erkennen, ob und wo gegebenenfalls „Schwachstellen“ im Unterricht vorkommen und bei welchem Schüler sich Lernschwierigkeiten entwickeln könnten. Dadurch ist es möglich, pädagogische Maßnahmen direkt oder zeitnah abzuleiten bzw. zu optimieren. Die formative Evaluation dient damit der Qualitätssicherung. Das Verfahren FE-RS 2 soll den genannten Merkmalen formativer Evaluation entsprechen und einen Beitrag zur veränderungssensitiven Diagnostik von Rechtschreibleistungen leisten. Die FE-RS 2 soll demnach ein formatives Verfahren zur Evaluation des Unterrichts sein. Vor dem Hintergrund der formativen Evaluation soll die FE-RS 2 ein veränderungssensitives Messinstrument sein, das empfindlich für Entwicklungsveränderungen im Lernbereich Rechtschreiben ist und Lernverläufe sichtbar macht. Da es Anwendung im regulären Unterricht finden soll, soll es vor dem Hintergrund von GOMs als Screeningverfahren entwickelt werden. Es soll nicht alle Facetten der Rechtschreibleistung wiedergeben, sondern einen Überblick über die Rechtschreibleistung der Schüler in ausgewählten Bereichen bieten. Mit dem Verfahren soll die Entwicklung der Rechtschreibleistung zeitlich engmaschig, mindestens im Abstand von 10 Wochen, dokumentiert werden. Damit sich Schüler nicht an die Aufgaben des Verfahrens erinnern und keine Lern- und Übungseffekte auftreten können, sollen zwei Parallelförmungen des Verfahrens konzipiert werden. Zudem soll angestrebt werden, dass das Verfahren so wenig Zeit wie möglich beansprucht und sich so besser in den Unterrichtsalltag integrieren lässt.

3 Theoretische Grundlagen des Rechtschreiberwerbs und curriculare Vorgaben für den Lernbereich Rechtschreiben

In diesem Kapitel werden die theoretischen Grundlagen des Rechtschreiberwerbs dargelegt, um das Verfahren zur Lernverlaufsdokumentation, die FE-RS 2, theoriegeleitet entwickeln zu können. Ausgehend von notwendigen Begriffsbestimmungen und linguistischen Grundlagen des Schriftspracherwerbs werden zuerst Modelle der Rechtschreibentwicklung und anschließend Modelle des Rechtschreibprozesses erläutert. Sodann wird analysiert, wie der Lernbereich Rechtschreiben in den Bildungsstandards der KMK (2005), in dem Rahmenplan für Mecklenburg-Vorpommern (MBJS et al., 2004) und den Sprachbüchern (MBWK, 2009a) verankert ist und ob er auf den aktuellen Theorien des Rechtschreiberwerbs basiert. Das Kapitel schließt mit einem Überblick über Möglichkeiten der Rechtschreibdiagnostik und Schlussfolgerungen für die Entwicklung der FE-RS 2.

3.1 Begriffsbestimmungen

Für die Auseinandersetzung mit Fragen des Rechtschreiberwerbs ist es notwendig, den hier zu betrachtenden Gegenstandsbereich des Rechtschreibens darzulegen. Die Bestimmung von Grundbegriffen ist dafür unerlässlich. Genaue begriffliche Unterscheidungen sind insofern erforderlich, da zum einen die linguistische Erforschung dieses Forschungsgegenstandes noch große Lücken aufweist (Nerius, 2007) und es zum anderen „ein heilloses Durcheinander“ (Ludwig, 1983, S. 1) von Begrifflichkeiten gibt. Begriffe unterschiedlicher Bedeutung wie z. B. Schriftlichkeit, geschriebene Sprache, Schriftsystem, Literalität, schriftliche Kommunikation, Schreiben, Schrift, Schriftsprache oder Mündlichkeit, gesprochene Sprache, Sprachsystem, Oralität, mündliche Kommunikation werden zum Teil synonym verwendet. Nicht nur die Verständlichkeit sondern vor allem das Nachvollziehen der Argumentationsstruktur ist dadurch erschwert (Ludwig, 1983). Deshalb werden im Folgenden die dieser Arbeit zugrunde liegenden Begriffe definiert und voneinander abgegrenzt.

Schriftsprache ist ein interdisziplinärer Forschungsgegenstand. Der Begriff Schriftsprache hat demnach mehrere Bedeutungsrichtungen. Nach Glück (2005) bezeichnet Schriftsprache:

1. das Vorhandensein technischer und linguistischer Grundlagen wie Zeichen, Regeln sowie Schreib- und Lesefähigkeiten für das Verschriften von Sprache,
2. die offiziell gültigen und verbindlichen Regelungen für die schriftliche Form von Sprache,
3. das selbstständige Aneignen der sprachlichen Strukturen und Funktionen beim Erwerb des Lesens und Schreibens.

Ein *Sprachsystem* beschreibt „die interne Ordnung sprachlicher Elemente (Phoneme, Morpheme, Sätze usw.) untereinander sowie ihren Funktionszusammenhang auf allen Beschreibungsebenen und in Relation zu sozialen, dialektalen u.a. [sic] Subsystemen“ (Bußmann, 2008, S. 662).

Ein *Schriftsystem* ist ein sprachabhängiges Inventar graphischer Zeichen. Es ist sprachabhängig, weil eine Schrift für die Verwendung in einer bestimmten Sprache nach sprachspezifischen Regeln verändert und erweitert worden ist (Ludwig, 1983). Der obigen Definition von Bußmann (2008) entsprechend, beschreibt ein Schriftsystem die interne Ordnung schriftsprachlicher Elemente untereinander.

Rechtschreibung (Orthographie) ist die Norm der geschriebenen Sprache. „Die Norm engt die in der geschriebenen Sprache vorkommenden und möglichen Varianten auf die allgemein anerkannten ein, sie umfaßt [sic] nur die zu einem bestimmten Zeitpunkt in der entsprechenden Gemeinschaft allgemein akzeptierten und überwiegend realisierten Varianten oder Möglichkeiten der Schreibung“ (Nerius, 2000, S. 28). Rechtschreibung reduziert alle möglichen Schreibweisen auf die durch die Norm festgelegten Schreibweisen, einschließlich der Zeichensetzung (Bußmann, 2008; Glück, 2005; Nerius, 2000).

Rechtschreiben ist ein kognitiv gesteuerter und wissensbasierter Prozess (Friedrich, 1996; Maas, 1991), die orthographisch zulässigen Schreibungen und Zeichensetzungen beim Schreiben zu wählen.

Eine Schulleistung, das heißt auch die *Rechtschreibleistung*, ist „das Ergebnis von Lernprozessen, die durch Unterrichtsmaßnahmen angeregt und/oder planvoll gesteuert wurden“ (Krapp, 1984, S. 46). Es ist zwischen der Schulleistung als singuläre

Einzelleistung, z. B. die Fehleranzahl in einem Diktat, und der Schulleistung als generalisierte Fähigkeit, z. B. die Fähigkeit, orthographisch richtig zu schreiben, zu unterscheiden (Krapp, 1984). Im Vergleich dazu wird der Begriff der *Rechtschreibkompetenz* verwendet, wenn anwendungsorientierte Kontexte der Leistung betont werden sollen. „Rechtschreibkompetenz ist [demnach] in ihrem Kern die Fähigkeit, . . . viele Wörter in verschiedenen Schreibsituationen richtig und weitgehend automatisiert schreiben zu können“ (Hinney, Huneke, A. Müller & Weinhold, 2008, S. 121f.).

3.2 Linguistische Grundlagen des Schriftspracherwerbs

Seit jüngster Vergangenheit findet die Schrift neben der Sprache als Gegenstand der germanistischen Linguistik wieder Beachtung. Folglich etabliert sie sich mit dem Forschungsfeld der *Schriftlinguistik* als eigenständiger Schwerpunkt der Linguistik (Nerius, 2007). Hinweise dafür sind entsprechende universitäre Lehrveranstaltungen, Tagungen und Publikationen, wie die von Dürscheid (2006) oder Neef und Weingarten (in Druck). Auch im „Metzler Lexikon Sprache“ (Glück, 2005) findet sich ein Eintrag zum Stichwort Schriftlinguistik: „Zusammenfassende Bez. [Bezeichnung] für Bemühungen, konsistente Beschreibungen und Analysen der . . . geschriebenen Sprachform von Spr. [Sprache] zu gewinnen und sie zu einer allgemeinen Schrifttheorie als konstitutivem Bestandteil einer allgemeinen Sprachtheorie zu entwickeln“ (S. 572).

Es war ein Anliegen der älteren Linguistik, die Frage nach dem Verhältnis zwischen dem Sprach- und dem Schriftsystem zu beantworten. Innerhalb dieser, in den 1980er Jahren im deutschsprachigen Raum (Enderle, 2005) geführten, Diskussion sind zwei Hauptpositionen zu unterscheiden: die *Autonomie-* und die *Abhängigkeitshypothese*. Vertreter der Autonomiehypothese postulieren die Unabhängigkeit des Schriftsystems vom Sprachsystem. Vertreter der Abhängigkeitshypothese hingegen nehmen an, dass das Schriftsystem aus dem Sprachsystem abgeleitet werden kann und damit kein unabhängiges System ist (Glück, 2005). Zudem führt Glück (2005) die *Interdependenzhypothese* an. Vertreter der Interdependenzhypothese fassen das Schriftsystem als eigenständiges System auf, bei dem die gesprochene Sprache Modell für die geschriebene ist. Die Interdependenzhypothese relativiert damit die traditionelle Autonomiehypothese und ist eine Variante dieser (Dürscheid, 2006). Die neuere Schriftlinguistik nimmt Abstand von

dieser Diskussion. Sie untersucht das Schriftsystem zum einen auf Strukturmerkmale des Sprachsystems und zum anderen auf eigene herausgebildete Strukturmerkmale (Weingarten, 2003).

In den folgenden Ausführungen werden die grundlegenden linguistischen Kennzeichen des deutschen Sprach- und Schriftsystems erläutert, um die Besonderheiten des Lerngegenstandes Schriftsprache und die damit verbundenen Schwierigkeiten beim Schriftspracherwerb zu verdeutlichen.

3.2.1 Aufbau des deutschen Sprach- und Schriftsystems

Das Sprachsystem besteht aus hierarchisch geordneten Ebenen. Jede Ebene besitzt eine *Konstruktionsfunktion* und ein *Konstruktionsmittel*. Die Konstruktionsfunktion beschreibt die Beziehung zwischen den Bestandteilen des Sprachsystems und der kommunikativen Funktion des Sprachsystems. Die sprachlichen Einheiten als Bestandteile des Sprachsystems dienen der Kommunikation, der Funktion von Sprache. Jede Einheit ist unterschiedlich weit von dieser Funktion entfernt und erfüllt sie direkt oder indirekt. Texte nehmen direkt, Einheiten niedriger Stufen indirekt an der Funktion teil. Letztere haben deshalb „innere Konstruktionsfunktionen“ (Daneš, 1982, S. 161). Konstruktionsmittel beschreiben sprachliche Zeichen, Einheiten und Merkmale von Ebenen, aus denen Einheiten der nächsthöheren Ebene konstruiert werden. Daraus ergibt sich, dass jede Ebene E_n sowohl Konstruktionsmittel der Ebene E_{n+1} als auch Funktion der Ebene E_{n-1} ist (Daneš).

Die Ebenen des Sprachsystems bestehen aus *unilateralen* oder *bilateralen Zeichen*. Dem Zeichenmodell von Saussure (2001) zufolge beziehen sich unilaterale Zeichen auf nur eine der beiden Seiten des sprachlichen Zeichens, die Formseite oder die Bedeutungsseite. Bilaterale Zeichen umfassen beide Seiten. Beide Seiten des sprachlichen Zeichens konstruieren Ebenen. Demnach gibt es bilaterale und unilaterale Ebenen. Zu den bilateralen Ebenen zählen die *Nullebene*, die *morphematische*, *lexikalische*, *syntaktische* und die *Textebene*. Zu den unilateralen Ebenen gehören die *Formebene*, die sich in die *phonologische* (E_p) und *graphische Ebene* unterteilt, und die *semantische Ebene* (Bedeutungsebene) (E_s) (Daneš, 1982). Daneš (1982) und Nerius (2007) verwenden den Begriff graphische Ebene. Dieser Terminologie wird in der vorliegenden Arbeit nicht

gefolgt, da die Graphematik die Grundeinheiten des Schriftsystems untersucht (Dürscheid, 2006). Damit dieser Bezug deutlich ist, wird die Bezeichnung *graphematische Ebene* (E_g) gewählt. Die unilateralen Ebenen konstruieren in Verbindung mit der Form- und Bedeutungsseite die bilateralen Ebenen des Sprachsystems (Nerius, 2007). Sie bilden gewissermaßen das Körpernetz, siehe Abbildung 5, der bilateralen Ebenen. Damit gehören beispielsweise zur phonologischen Ebene nicht nur Erscheinungen, die das Phonem betreffen, sondern auch die phonologische Seite aller bilateralen Ebenen. Zur Veranschaulichung ist der Aufbau des Sprachsystems in Abbildung 5 dargestellt.

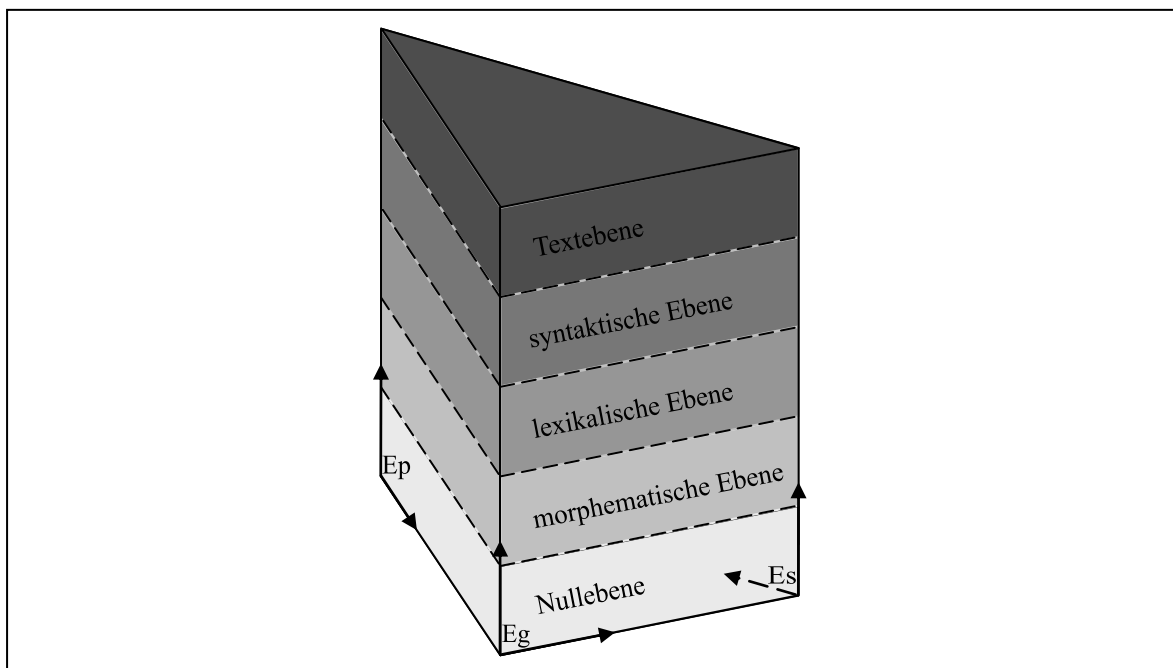


Abbildung 5: Der Aufbau des Sprachsystems

Anmerkungen: E_p = phonologische Ebene, E_g = graphematische Ebene, E_s = semantische Ebene

3.2.1.1 Sprachliche Einheiten der unilateralen Ebenen

Die sprachlichen Einheiten der unilateralen Ebenen haben eine zeichenbildende Funktion (Daneš, 1982). Sie bilden die sprachlichen Zeichen der bilateralen Ebenen, welche aus einer festen Verbindung von Form- (Zeichenkörper, das heißt ob lautlich oder graphisch realisiert) und Bedeutungsseite (Zeicheninhalt) bestehen (Saussure, 2001).

Das kleinste Element der phonologischen Ebene ist das *distinktive phonologische Merkmal*. Bündel distinktiver phonologischer Merkmale bilden *Phoneme*. Phoneme selbst

tragen keine Bedeutung, sondern sind „kleinste bedeutungsunterscheidende Einheiten“ (Nerius, 2007, S. 77). Durch die Bildung von Minimalpaaren lassen sich Phoneme nachweisen. Distinktive phonologische Merkmale sind bedeutungsunterscheidende Lauteigenschaften des Phonems. Sie charakterisieren und unterscheiden Phoneme voneinander. Phoneme können sich bestimmte Merkmale mit anderen Phonemen teilen, unterscheiden sich aber mindestens in einem Merkmal von allen anderen (Grassegger, 2010). Zu den distinktiven phonologischen Merkmalen zählen nach Grassegger (2010):

1. *Prosodische Merkmale*, die sich über größere sprachliche Einheiten wie z. B. die Silbe erstrecken und nur in Relation zu anderen sprachlichen Einheiten definiert werden können. Zu den prosodischen Merkmalen zählen die *Quantität* (Lautdauer, z. B. /mi:tə/ *Miete* vs. /mitə/ *Mitte*), der *Akzent* (Betonung, z. B. /'kafə/ *Kaffee* vs. /ka'fe:/ *Café*) und die *Intonation* (Tonhöhenbewegung, z. B. /komən ↓zi:/ *Kommen Sie!* vs. /komən ↑zi:/ *Kommen Sie?*),
2. *Segmentale Merkmale*, die an einzelne Phoneme gebunden sind und unabhängig von anderen sprachlichen Einheiten definiert werden können. Zu den segmentalen Merkmalen zählen die *Oberklassenmerkmale* (Lautklassenzuordnung, z. B. /s/ [+konsonantisch], [-sonorant], [+obstruent]), der *Artikulationsort* (Zungenstellung, z. B. /s/ [-hoch], [-hinten], [+koronal], [+anterior]) und die *Artikulationsart* (Luftstrombehinderung, z. B. /s/ [+kontinuierlich], [-nasal], [-lateral], [-stimmhaft]).

Die nächsthöhere Einheit der phonologischen Ebene ist die *Silbe*. Sie ist die kleinste Phonemsequenz und regelt die artikulatorische Gliederung des Lautstroms (Nerius, 2007). Nach phonologischen Kriterien bestehen Silben aus einem vokalischen Silbenkern und einer konsonantischen Silbenschale, die auch fehlen kann. Die Silbenschale besteht aus einem Silbenkopf (Anfangsrand) und einer Silbenkoda (Endrand). Silbenkern und Silbenkoda bilden zusammen den Reim (Grassegger, 2010). Am Beispiel des Wortes <Bein> mit der einfachen Silbenstruktur Konsonant (K), Vokal (V), Konsonant (K) wird in Abbildung 6 der Aufbau einer Silbe dargestellt.

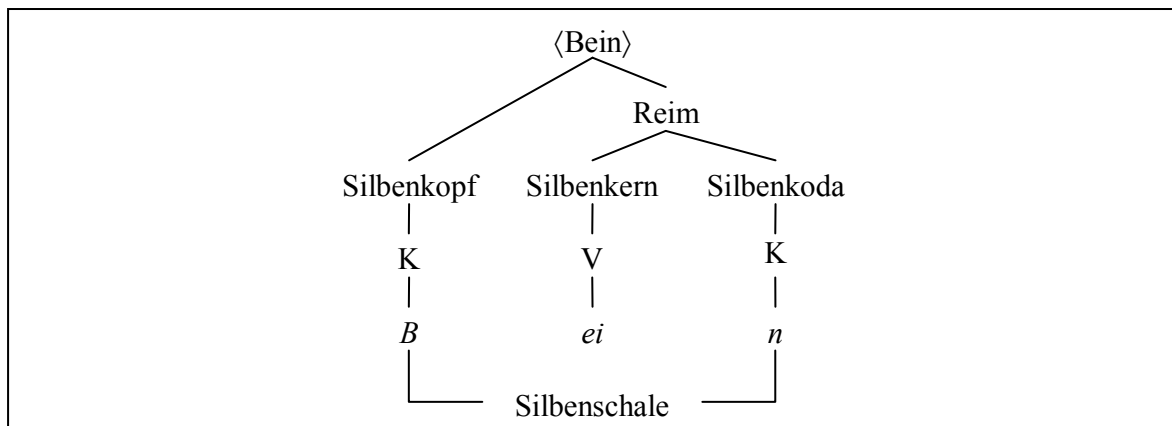


Abbildung 6: Der Silbenaufbau am Beispiel *Bein* (in Anlehnung an Grassegger, 2010)
 Anmerkungen: K = Konsonant, V = Vokal

Die nächsthöheren Einheiten der phonologischen Ebene sind das *phonologische Morphem* und das *phonologische Wort*, die mit Silben übereinstimmen können. Weitere Einheiten sind der *phonologische Satz*, bei dem phonologische Wörter zu Wortgruppen und Sätzen zusammengefasst werden, und der *phonologische Text* (Nerius, 2007). Bei Nerius (2007) ist sowohl von phonischen Morphem-, Wort-, Satz- und Textformen als auch von phonologischen Wörtern zu lesen, wobei eine Bedeutungsunterscheidung nicht erkennbar ist. Entsprechend der phonologischen Formseite sprachlicher Zeichen wird in der vorliegenden wie bei Eisenberg (1996) ausschließlich vom phonologischen Wort, analog dazu vom graphematischen Wort etc., gesprochen.

Die kleinste Einheit der semantischen Ebene ist das *Sem*. Diese „kleinsten distinktiven Bedeutungskomponenten“ (Bußmann, 2008, S. 614) beschreiben die Bedeutung eines sprachlichen Ausdrucks. Semantische Merkmale dienen der Unterscheidung von Wortbedeutungen. „Sie bilden nicht unmittelbar physikalische Eigenschaften der realen Welt ab, sondern spiegeln die psychischen Bedingungen, gemäß denen die Umwelt durch den Menschen bzw. dessen Sprache verarbeitet wird“ (Bußmann, 2008, S. 618). Das Wort *Stift* z. B. bezeichnet unterschiedliche Sachverhalte und wird semantisch differenziert: *Stift*₁ (Schreibgerät), *Stift*₂ (Altersheim) und *Stift*₃ (halbwüchsiger Junge). Seme verbinden sich zu *Semenen* und beschreiben die Gesamtbedeutung von Morphemen, Wörtern, Sätzen und Texten (Nerius, 2007).

Die kleinste Einheit der graphematischen Ebene ist das *distinktive graphematische Merkmal*, z. B. ein Strich oder Kreis (Nerius, 2007). Bündel distinktiver graphematischer Merkmale bilden *Grapheme*. Grapheme sind analog zum Phonem „kleinste distinktive

Einheiten eines Schriftsystems“ (Bußmann, 2008, S. 246). Nerius (2007) verwendet die Begriffe Graphem und *Buchstabe* synonym. Den Graphemen ordnet er Buchstaben und Nichtbuchstaben wie Zahlen, Satzzeichen und Symbole zu. In der vorliegenden Arbeit hingegen wird zwischen Graphemen und Buchstaben unterschieden. Es wird davon ausgegangen, dass zur Klasse der Grapheme nur Buchstaben zählen und nicht jedem Buchstaben ein Graphem (und umgekehrt) zugeordnet werden kann (Dürscheid, 2006; Eisenberg, 2006). Ob es sich bei Buchstaben und Buchstabenkombinationen um Grapheme handelt, lässt sich durch die Bildung von Minimalpaaren nachweisen. Wörter, die sich in gleicher Position unterscheiden und durch den Austausch dieses einen Buchstabens in ihrer Bedeutung ändern, bilden ein Minimalpaar, z. B. laufen/kaufen, mein/dein. Damit sind *l*, *k*, *m* und *d* Grapheme. Bei den Wörtern *Chor*, *Qualle* und *Schaf* lassen sich jedoch nur die ersten zwei bzw. drei Buchstaben gemeinsam ersetzen. Damit sind auch *ch*, *qu* und *sch* Grapheme (Eisenberg, 2006). Nur Buchstaben, die bedeutungsunterscheidend sind, bilden Grapheme und zählen zum Grapheminventar (Dürscheid, 2006). Nach Eisenberg (2006) zählen zum Grapheminventar des Deutschen die Konsonantengrapheme *p*, *t*, *k*, *b*, *d*, *g*, *f*, *s*, *β*, *w*, *j*, *h*, *m*, *n*, *l*, *r*, *qu*, *ch*, *sch*, *z* und die Vokalgrapheme *a*, *e*, *i*, *ie*, *o*, *u*, *ä*, *ö*, *ü*. Die Buchstaben *c*, *v*, *x*, *y* kommen nicht im Grapheminventar vor. Sie werden im Deutschen nur bei Fremdwörtern wie *Clown* und *Hobby* verwendet oder kommen wie das *v* und *x* in *Vogel* und *Hexe* im Kernbestand seltener vor. Sie gehören deshalb nicht zum Kernbestand des Grapheminventars (Eisenberg, 2006). Die nächsthöhere Einheit der graphematischen Ebene ist das *graphematische Wortsegment*, welches bei Worttrennungen am Zeilenende auftritt und auf phonologischer Ebene mit der Silbe übereinstimmen kann. Weitere Einheiten sind das *graphematische Morphem* und das *graphematische Wort* als Graphemfolge mit Spatien (Leerstellen). Die höchsten Einheiten der graphematischen Ebene sind der *graphematische Satz* mit Interpunktionen und der *graphematische Text* mit den dazugehörigen Kennzeichen wie Überschriften, Absätzen und Textgliederungszeichen (Nerius, 2007).

3.2.1.2 Sprachliche Einheiten der bilateralen Ebenen

Die sprachlichen Einheiten der bilateralen Ebenen haben eine Konstruktionsfunktion. Sie werden, bis auf einige Ausnahmen, aus Einheiten der nächstniedrigeren Ebene konstruiert. Somit sind niedrige Einheiten in höheren enthalten. Unterhalb aller bilateralen Ebenen

befindet sich die Nullebene. Sie enthält keine bilateralen Zeichen, sondern Konstruktionsmittel für die morphematische Ebene. Zu diesen Konstruktionsmitteln zählen die Grundbausteine der unilateralen Ebenen: die phonologischen distinktiven Merkmale und Phoneme, die graphematischen distinktiven Merkmale und Grapheme sowie die Seme (Daneš, 1982; Nerius, 2007).

Die niedrigste bilaterale Ebene ist die morphematische Ebene. Die sprachliche Einheit dieser Ebene ist das *Morphem*, „die kleinste bedeutungstragende Einheit der Sprache“ (Nerius, 2007, S. 79). Nach Gallmann (1985) lassen sich drei Arten von Morphemen unterscheiden:

1. *Grundmorpheme*, die allein oder in Verbindung mit mehreren Grundmorphemen und Flexionsmorphemen Wörter bilden können, z. B. *Saft*,
2. *Wortbildungsmorpheme*, die nur in Verbindung mit Grundmorphemen Wörter bilden können, z. B. *saft-ig*,
3. *Flexionsmorpheme*, die keine wortbildende, sondern eine grammatische Funktion haben und nur in Verbindung mit Grund- und Wortbildungsmorphemen Wörter bilden können, z. B. *saftig-er*.

Als bilaterale Zeichen haben Morpheme eine Form- und Bedeutungsseite. Das phonologische Morphem kann der Silbe, das graphematische Morphem dem Wortsegment entsprechen. In Bezug auf die nächsthöhere Ebene haben Morpheme eine wortbildende Funktion (Nerius, 2007).

Die nächsthöhere bilaterale Ebene ist die lexikalische Ebene. Die sprachlichen Einheiten dieser Ebene sind das *Wort* und das *Lexem* (Nerius, 2007). Das Wort ist eine schwer zu definierende sprachliche Einheit. Grundlegend ist zwischen dem *grammatischen* und dem *lexikalischen Wort* zu unterscheiden. Das grammatische Wort wird nach grammatischen Merkmalen wie Kasus, Numerus und Tempus flektiert. Das lexikalische Wort, auch Lexem genannt, ist eine „abstrakte Basiseinheit des ... Lexikons“ (Glück, 2005, S. 379) und nicht durch Flexion gekennzeichnet (Glück, 2005). So handelt es sich bei den Beispielen *des Saftes* (Genitiv, Singular) und *die Säfte* (Nominativ, Plural) um zwei grammatische Wörter desselben Lexems *Saft*. Als bilaterale Zeichen haben Wörter eine Form- und eine Bedeutungsseite. In Bezug auf die nächsthöhere Ebene haben Wörter und Lexeme eine wortgruppen- und satzbildende Funktion (Nerius, 2007). Wird im Folgenden von Wörtern gesprochen, ist damit das grammatische Wort gemeint.

Die nächsthöhere bilaterale Ebene ist die syntaktische Ebene. Die sprachlichen Einheiten dieser Ebene sind die *Wortgruppe* und der *Satz* (Nerius, 2007). Eine Wortgruppe „ist nicht nur eine Abfolge von Wörtern, sondern eine Gruppe von syntakt. [syntaktisch] zusammengehörenden Wörtern“ (Glück, 2005, S. 669), z. B. ein saftiger Apfel. Der Satz ist ein viel diskutierter Begriff ohne eine bislang allgemeingültige Definition. Ohne Anspruch auf Vollständigkeit und Richtigkeit soll hier ein Satz als sprachliche Einheit verstanden werden, die aus einem Wort oder mehreren Wörtern besteht. Als bilaterale Zeichen haben Wortgruppen und Sätze eine Form- und eine Bedeutungsseite. Zu den phonologischen Formen zählen hier der Satzakzent sowie die Tonhöhenbewegungen und Pausen. Zu den graphematischen Formen gehören die Interpunktionszeichen und die Großschreibung. Den syntaktischen Bedeutungen können z. B. Frage, Aussage und Aufforderung zugeordnet werden. In Bezug auf die nächsthöhere Ebene haben Wortgruppen und Sätze eine textbildende Funktion (Nerius, 2007).

Die höchste bilaterale Ebene ist die Textebene. Die sprachlichen Einheiten dieser Ebene sind die *Textzeichen*. Textzeichen bestehen aus Sätzen, die zueinander in Beziehung gesetzt werden. Als bilaterale Zeichen haben Texte eine Form- und eine Bedeutungsseite (Nerius, 2007).

Die vorangegangenen Ausführungen bekräftigen, dass das Schriftsystem Teil des Sprachsystems ist. Das Schriftsystem ist kein gesonderter Bereich, sondern steht mit der graphematischen Ebene in enger Beziehung zu allen anderen Ebenen des Sprachsystems. Die graphematische Ebene ist durch Einheiten anderer Ebenen gekennzeichnet, z. B. die sprachlichen Zeichen der bilateralen Ebenen. Gleichzeitig besitzt sie aber auch Einheiten, die nur ihr eigen sind, z. B. die Grapheme. Demzufolge besitzt die graphematische Ebene wie alle Ebenen innerhalb des Sprachsystems eine relative Autonomie. Für den Erwerb der Schriftsprache bedeutet das, dass Schüler Einsichten in die Struktur und den Form-Funktions-Zusammenhang der sprachlichen Einheiten gewinnen sollten. So müssen sie z. B. die Bedeutung eines Wortes erkennen, um es in Morpheme gliedern und Wortfamilien bilden zu können. Die Einsichten in sprachliche Einheiten entwickeln sich jedoch nicht durch die formale linguistische Sprachbetrachtung, sondern durch das Operieren mit sprachlichen Einheiten (Nerius, 2007).

3.2.2 Prinzipien und Regeln der konventionellen Orthographie

Die Kernfunktion von Schriftsystemen ist das Rekodieren, das Umwandeln von geschriebener in gesprochene Sprache. Die Aufgabe von Schriftsystemen ist demnach, jeder geschriebenen Form durch Korrespondenzregeln eine phonologische Repräsentation zuzuordnen. Der Grundgedanke dabei ist, die phonologischen Repräsentationen in Schreibungen erkennbar zu machen. Für diese Aufgabe gibt es innerhalb des Schriftsystems zwei Teilbereiche, die *Graphematik* und die *Orthographie* (Neef, 2005b). Die Graphematik ist der Teilbereich des Schriftsystems, der die graphematischen und phonologischen Grundeinheiten sowie ihre Beziehungen untereinander beschreibt (Bußmann, 2008; Glück, 2005; Neef, 2005b). Sie gilt als ein „natürliches System sprachlichen Wissens“ (Neef, 2005b, S. 8) und betrachtet die Menge möglicher Schreibungen für eine phonologische Repräsentation. Ihr gegenüber steht die Orthographie als „ein explizit geregeltes konventionelles System“ (Nerius, 2005b, S. 8). Sie wählt aus der Menge möglicher Schreibungen ein Wort aus und erklärt es als orthographisch korrekt. Demnach gibt es viele, prinzipiell unbegrenzte, mögliche Schreibungen zu phonologischen Repräsentationen, aber für gewöhnlich nur eine orthographisch richtige. Beispielsweise gibt es für das Wort *Wal* sechs graphematisch mögliche Schreibungen: ⟨Val⟩, ⟨Vaal⟩, ⟨Vahl⟩, ⟨Wal⟩, ⟨Waal⟩ und ⟨Wahl⟩. Aber nur ⟨Wal⟩ ist für den Meeressäuger die orthographisch zulässige Form. Enthält der graphematische Lösungsraum nur ein Element, ist es wahrscheinlich die orthographisch korrekte Schreibung für die jeweilige phonologische Repräsentation. In der Regel enthält der graphematische Lösungsraum aber mehr als ein Element, sodass die orthographische Komponente des Schriftsystems die orthographisch korrekte Schreibung eines Wortes auswählen muss. Der orthographischen Beschränkung gelingt es aber zumeist nicht, die orthographisch richtige Schreibung eines Wortes, das heißt eine einzige Schreibung, vorherzusagen. Aufgrund dessen muss zwischen einer *systematischen* und *konventionellen Orthographie* unterschieden werden. Die systematische Orthographie begrenzt durch orthographischer Beschränkungen die möglichen Schreibungen für ein Wort, um den graphematischen Lösungsraum sukzessive einzugrenzen. Die konventionelle Orthographie bestimmt die einzig orthographisch korrekte Schreibung eines Wortes (Neef, 2005b). Sie entspricht damit den amtlichen Regelungen des Rats für deutsche Rechtschreibung (2006). Vorwiegend stützt sich die konventionelle auf die systematische Orthographie. Sie kann aber auch Schreibungen für richtig erklären, die nicht im graphematischen Lösungsraum enthalten sind, z. B. ⟨doofe⟩

für [do.və] (Neef).

Die konventionelle Orthographie verfährt nach *Prinzipien* und *Regeln*, um die einzig orthographisch korrekte Schreibung eines Wortes zu bestimmen. Sie erfasst den amtlich festgelegten Satz von Schreibungen (Neef, 2005b). Aufgrund der Alphabetschrift, wie sie das Deutsche verwendet, und ihrer Korrespondenz zwischen Phonemen und Graphemen beruht die konventionelle Orthographie auf zwei Grundprinzipien: (1) dem *phonologischen Grundprinzip*, welches die Beziehung zwischen phonologischer und graphematischer Ebene beschreibt und (2) dem *semantischen Grundprinzip*, welches die Beziehung zwischen phonologischer und semantischer Ebene beschreibt (Nerius, 2007; Rat für deutsche Rechtschreibung, 2006). Die beiden Grundprinzipien bestehen aus weiteren Prinzipien, die die Beziehungen der phonologischen, graphematischen und semantischen Ebene untereinander genauer bestimmen. Dem phonologischen Grundprinzip wird

1. das *phonematische Prinzip*, das die Korrespondenz von Phonemen und Graphemen betrifft, z. B. /li:bən/ und ⟨lieben⟩,
2. das *syllabische Prinzip*, das die Beziehung zwischen der Silbe und dem graphematischen Wortsegment beschreibt, z. B. /kraft.vɛrk/ und ⟨Kraft-werk⟩,
3. das *intonatorische Prinzip*, das die Abbildung prosodischer Einheiten auf der graphematischen Ebene, z. B. durch Satzzeichen, betrifft, zugeordnet.

Dem semantischen Grundprinzip wird

4. das *morphematische Prinzip*, das die Konstanzschreibung von Morphemen fordert, z. B. Nummer → nummerieren,
5. das *lexikalische Prinzip*, das bestimmte semantische Eigenschaften durch verschiedene graphematische Mittel erfasst, z. B. Großschreibung der Nomen,
6. das *syntaktische Prinzip*, das Satzbedeutungen durch graphematische Mittel erfasst, z. B. Großschreibung am Satzanfang und
7. das *textuale Prinzip (Textprinzip)*, das Aspekte der Textbedeutung durch graphematische Mittel betont, z. B. Überschriften und Absätze, zugeordnet (Nerius, 2007).

Auf der Grundlage der Prinzipien als eigentliche Basis für orthographisch richtige Schreibungen werden Regeln bestimmt. Sie legen fest, wie Prinzipien angewendet werden und führen zu eindeutigen Schreibungen (Gallmann & Sitta, 1998). Zu den Regeln zählen die Laut-Buchstaben-Zuordnung, die Getrennt- und Zusammenschreibung, die Schreibung mit Bindestrich, die Groß- und Kleinschreibung, die Zeichensetzung und die Worttrennung

am Zeilenende (Rat für deutsche Rechtschreibung, 2006). Insbesondere in den Bereichen der Laut-Buchstaben-Zuordnung, der Getrennt- und Zusammenschreibung sowie in wenigen Fällen der Groß- und Kleinschreibung konnte durch die Regeln aber keine eindeutige Handlungsanweisung erreicht werden (Gallmann & Sitta, 1998). Erstgenanntes ist mit der Alphabetschrift des Deutschen zu begründen, die vorrangig lautliche Eigenschaften der gesprochenen Sprache repräsentiert, denn dem Alphabet, als begrenztes Inventar von Graphemen, steht eine größere Menge von phonetischen Realisierungsmöglichkeiten gegenüber. Die Zuordnung zwischen Lautung und Schrift erfolgt auf der Ebene der Phoneme und Grapheme, weil sie Träger distinktiver Merkmale sind. Zwar gibt es auf der Zuordnungsebene zwischen Phonemen und Graphemen viele eindeutige Zuordnungen, aber keine durchgehenden Korrespondenzen. So entsprechen z. B. dem Phonem /k/ mit ⟨k⟩ in *Kohl*, ⟨kk⟩ in *Sakko*, ⟨ck⟩ in *Acker*, ⟨g⟩ in *Tag*, ⟨gg⟩ in *flaggt*, ⟨ch⟩ in *Fuchs* und ⟨c⟩ in *Clown* verschiedene Schreibungen. Dem Graphem ⟨ch⟩ z. B. entsprechen mit [x] in *Dach*, [ç] in *China*, [k] in *Dachs*, [ʃ] in *Charme* und [tʃ] in *Chip* verschiedenen Lautungen (Ramers, 2001). Infolgedessen müssen Schreibungen in den genannten Bereichen von *Einzelfestlegungen* bestimmt werden (Gallmann & Sitta, 1998).

Es zeigt sich, dass der graphematische Lösungsraum eine prinzipiell unbegrenzte Anzahl von Schreibungen für phonologische Repräsentationen bereithält. Für die Graphematik als einen Teilbereich des Schriftsystems ist dies kein Problem, wohl aber für die Orthographie. Da trotz orthographischer Beschränkungen die orthographisch lizenzierte Schreibung für eine bestimmte phonologische Repräsentation nicht vorhersagbar ist, muss zwischen einer systematischen und einer konventionellen Orthographie unterschieden werden. Die systematische Orthographie begrenzt den graphematischen Lösungsraum auf ein Minimum möglicher Schreibungen. Die konventionelle Orthographie legt die orthographisch richtige Schreibung auf der Grundlage von Prinzipien, Regeln und Einzelfestlegungen fest.

3.2.3 Zusammenfassung und Schlussfolgerungen

Wie in diesem Unterkapitel herausgearbeitet, ist das Schriftsystem mit seiner graphematischen Ebene fester Bestandteil des Sprachsystems. Das Sprachsystem besteht aus hierarchisch geordneten Ebenen, den unilateralen und bilateralen Ebenen. Die unilateralen Ebenen haben eine zeichenbildende, die bilateralen Ebenen eine

Konstruktionsfunktion. Innerhalb des Schriftsystems ist nach Neef (2005b) zwischen Graphematik und Orthographie zu unterscheiden. Beide Teilbereiche des Schriftsystems dienen der Kernfunktion von Schriftsystemen, dem Rekodieren. Gilt die Graphematik als ein natürliches System möglicher Schreibungen, steht ihr die Orthographie als ein konventionelles System gegenüber. Die Orthographie verfolgt das Ziel, die tatsächlichen Schreibungen von Wörtern vorherzusagen. Das gelingt jedoch nicht immer. Für den Lerngegenstand Schriftsprache, insbesondere den Rechtschreiberwerb, hat dies zur Folge, dass „viel orthographisches Wissen einfach auswendig gelernt werden [muss]“ (Neef, 2005b, S. 204). Diese Tatsache betont, „dass die Sachstruktur [der Orthographie] nicht der Lernstruktur entspricht. Kinder müssen sich das System der Schrift selbst konstruieren“ (Nickel, 2006, S. 70). In einem ersten Schritt müssen sie sich die Korrespondenzen zwischen Phonemen und Graphemen, das heißt *graphematisches Können*, aneignen. In einem zweiten Schritt müssen Kinder lernen, die orthographisch richtige Schreibung zu verwenden, das heißt sich *orthographisches Können* anzueignen (Neef, 2005a). Wie Kinder in diesem Sinne Rechtschreiben lernen, wird nachfolgend dargelegt. Auf der Basis der in diesem Abschnitt dargestellten linguistischen Grundlagen wird in Abschnitt 3.3 ein Überblick über Modelle der Rechtschreibentwicklung gegeben.

3.3 Modelle der Rechtschreibentwicklung

Der Rechtschreiberwerb ist ein Entwicklungs- und Lernprozess, der sich über die gesamte Schulzeit erstreckt. Er basiert auf den Zugängen der Kinder zur Sprache, ihren Einsichten in den Aufbau und die Funktion von Schrift sowie auf strukturierten Lehr- und Lernprozessen. Die Rechtschreibentwicklung erfolgt nicht gleichmäßig und kontinuierlich, sondern eher nach und nach (Scheerer-Neumann, 2008). Die Annahme, dass es sich hierbei um eine strikte Stufenabfolge mit klar voneinander abgrenzbaren Phasen wie in älteren Entwicklungsmodellen von Frith (1985, 1986) und K. B. Günther (1986) handelt, wird in der aktuellen Forschungsliteratur abgelehnt (Eichler, 1986; G. Thomé, 2006). Dennoch beeinflusste das Entwicklungsmodell von Frith die deutsche Forschung zum Schriftspracherwerb maßgeblich und bildete als Rahmenmodell die Grundlage für weitere Entwicklungsmodelle (Scheerer-Neumann, 1998; Stock & Schneider, 2008b). In der aktuellen Forschungsliteratur zur Rechtschreibentwicklung setzt sich zunehmend das Konzept der sich nach und nach entfaltenden und parallel verfügbaren

Rechtschreibstrategien durch (Böhme & Bremerich-Vos, 2009). Dieser Annahme entsprechen neuere Entwicklungsmodelle wie die von Löffler und Meyer-Schepers (2005, 2007, 2008), May (2002, 2008) und Scheerer-Neumann (2007, 2008).

In den folgenden Ausführungen werden die Entwicklungsmodelle von Frith (1985, 1986), dessen Erweiterung von K. B. Günther (1986) sowie die Modelle von Löffler und Meyer-Schepers (2005, 2007, 2008), May (2002, 2008) und Scheerer-Neumann (2007, 2008) erläutert, um die idealtypischen Stufen bzw. Phasen und Strategien, die Kinder bei der Rechtschreibentwicklung durchlaufen, zu verdeutlichen.

3.3.1 Das Sechsstufenmodell von Frith (1985, 1986) und das Fünf-Phasen-Modell von K. B. Günther (1986)

Das Entwicklungsmodell von Frith (1985, 1986) ist in seiner ursprünglichen Konzeption ein *Dreiphasenmodell* für den Leseerwerb. Das Dreiphasenmodell basiert auf der Annahme, dass sich die Leseentwicklung in drei aufeinander aufbauenden Phasen vollzieht. Jede Phase verweist auf die in der jeweiligen Phase dominierende Strategie, die *logographemische*, *alphabetische* oder *orthographische Strategie*. Die logographemische Strategie dominiert die erste Phase der Leseentwicklung. Die zweite Phase wird von der alphabetischen, die dritte Phase des Lesenlernens von der orthographischen Strategie angeführt. Der Wechsel von einer Strategie zur nächsten erfolgt, wenn sich Elemente der vorherigen Strategie mit Elementen der darauf aufbauenden Strategie verbinden und ineinander übergehen. Dabei wird die vorherige Strategie nicht aufgegeben, sondern in der neuen Strategie weitergeführt.

In der Erweiterung des Dreiphasenmodells der Leseentwicklung hin zu einem *Sechsstufenmodell* des Schriftspracherwerbs wird Frith (1985, 1986) den sich gegenseitig bedingenden Komponenten des Lesens und Rechtschreibens gerecht. Das Sechsstufenmodell besteht aus den bereits bekannten drei aufeinander aufbauenden Phasen. Die Phasen der Rechtschreibentwicklung sind in ähnlicher Weise charakterisiert wie die der Leseentwicklung. Neu ist, dass jede Phase in zwei Stufen, a und b, unterteilt ist, die das Lesen und Rechtschreiben betreffen. Die Stufen zeigen an, welche der beiden Komponenten die neue Strategie einleitet und wie stark die jeweiligen Strategien bereits

ausgeprägt sind. Bei einem Niveau von 1 sind grundlegende, bei 2 fortgeschrittene und bei 3 voll entfaltete Strategieformen anzunehmen. Ab Niveau 2 wird die jeweilige Strategie vom Lesen auf das Rechtschreiben bzw. umgekehrt übertragen. Das Sechsstufenmodell von Frith (1985) ist in Abbildung 7 dargestellt.

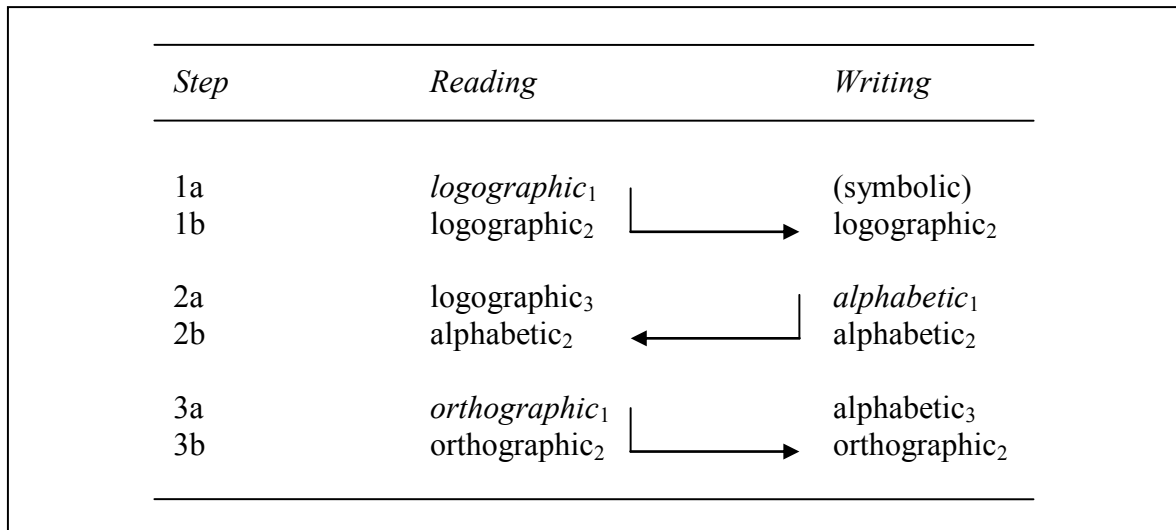


Abbildung 7: The Six-step Model of Skills in Reading and Writing Acquisition (Frith, 1985, S. 311)

Die erste Phase, die logographemische, beruht auf den „symbolic skills“ (Frith, 1985, S. 308) des Kindes, die es sich in einer präliteral-symbolischen Phase angeeignet hat. In der logographemischen Phase orientiert sich das Kind an hervorstechenden graphischen Merkmalen eines Wortes. Von diesen Merkmalen schließt es auf das ihm bekannte Wort. Dabei spielen die Graphemreihenfolge und die phonologischen Merkmale keine Rolle. Wörter ohne optische Schlüsselreize werden erraten. Die logographemische Strategie wird zuerst für das Lesen verwendet (Frith, 1985, 1986), da die Wahrnehmung der Schrift ihrer Erzeugung durch das Kind vorausgehen muss (Nickel, 2006). Erst wenn das Niveau 2 dieser Strategie erreicht ist und das Kind unmittelbar Wörter und kurze Sätze erkennt, kann die logographemische Strategie auch für das Schreiben genutzt werden (Frith, 1985, 1986). Charakteristisch für diese Phase ist, dass Kinder häufig Grapheme vertauschen und ihre Schriftproduktionen nicht lesen können. Grund dafür ist, dass sie in der logographemischen Phase nur einen visuellen Zugriff auf die Schrift haben und ihnen keine phonologische Strategie zur Verfügung steht (Nickel, 2006).

Der Wechsel von der logographemischen zur alphabetischen Phase erfolgt, wenn das Kind erste Phonem-Graphem-Korrespondenzen ergänzend zur logographemischen Strategie

nutzt. In der zweiten Phase, der alphabetischen, erkennt das Kind den Lautbezug der Schrift. Daran schließt sich die vollständige Aneignung der Phonem-Graphem-Korrespondenzen an. Die alphabetische Strategie wird zuerst für das Schreiben verwendet, wobei jedem Phonem ein Graphem zugeordnet wird. Erst wenn das Niveau 2 dieser Strategie erreicht ist und das Kind auch unbekannte Wörter lautgetreu verschriftet, kann die alphabetische Strategie auch für das Lesen genutzt werden. Das Lesen von Wörtern erfolgt durch graphemweises Rekodieren (Frith, 1985, 1986). Häufig kommt es in dieser Phase vor, dass Kinder Schwierigkeiten in der Entwicklung der alphabetischen Strategie haben und zu Strategiemischformen greifen. Sie versuchen, durch graphische Wortmerkmale in Kombination mit ersten Graphem-Phonem-Korrespondenzen auf das ganze Wort zu schließen oder sie versuchen, sich Wortschreibungen zusätzlich über den visuellen Zugriff zu merken (Nickel, 2006).

Der Wechsel von der alphabetischen zur orthographischen Phase erfolgt, wenn das Kind erste größere sprachliche Einheiten ergänzend zur alphabetischen Strategie nutzt. In der dritten Phase, der orthographischen, orientiert sich das Kind an größeren sprachlichen Einheiten wie Silben und Morphemen. Wörter werden durch das Wissen über die Struktur der Schriftsprache zum Teil direkt erkannt. Die orthographische Strategie wird zuerst für das Lesen verwendet, wodurch ein schnelleres und flüssigeres Lesen möglich ist. Erst wenn das Niveau 2 dieser Strategie erreicht ist, kann die orthographische Strategie auch für das Schreiben genutzt werden. Dann ist ein orthographisch richtiges Schreiben möglich (Frith, 1985, 1986). Auch in dieser Phase kann es zu Strategiemischformen kommen, sodass Kinder versuchen, durch Hinzunahme der alphabetischen Strategie alle orthographischen Phänomene zu erhören (Nickel, 2006). Obwohl mit der dritten Phase die Schriftsprachentwicklung in ihren Grundzügen als abgeschlossen gelten kann, ist eine sich anschließende vierte Phase anzunehmen. „However, at least a *Phase 4* is conceivable where the independence of written and spoken language is achieved and written language is handled as a system in its own right“ (Frith, 1985, S. 309).

K. B. Günther (1986) adaptierte das Sechsstufenmodell von Frith (1985, 1986) für den deutschen Sprachraum. Er erweiterte das Entwicklungsmodell um zwei Phasen, die bei Frith bereits angedeutet waren, die *präliteral-symbolische* und die *integrativ-automatisierte Phase*. Das Stufenmodell besteht demnach aus fünf Phasen, die jeweils in zwei Stufen unterteilt sind. Die einzelnen Phasen verweisen auf die jeweils dominierenden

Strategien und bauen aufeinander auf. Die Strategien bleiben in nächsthöheren Phasen latent vorhanden (K. B. Günther, 1986).

K. B. Günther (1986) stellt der logographemischen Phase die präliteral-symbolische Phase voran. Die präliteral-symbolische Phase basiert auf der Annahme, dass der Schriftspracherwerb nicht erst mit dem eigentlichen Erwerb des Lesens und Schreibens beginnt, sondern bereits vorher. Für K. B. Günther (1986) setzt diese Phase ein, wenn Kinder eine Beziehung zwischen dreidimensionalen Körpern und zweidimensionalen Bilddarstellungen herstellen können, weil das Erkennen des im Bild Dargestellten ein höheres Maß an Abstraktionsfähigkeit als die vorherigen sensomotorischen Leistungen vom Kind verlangt. Dabei handelt es sich um eine Vorstufe des Lesens. Das präliteral-symbolische Schreiben, das heißt das kindliche Zeichnen, setzt verspätet ein. Kinder zeichnen in dieser Phase für sie bedeutungsvolle Gegenstände, ohne den Anspruch auf Detailgenauigkeit und Vollständigkeit zu verfolgen. Ebenso ahmen sie das Schreiben nach. Die präliteral-symbolische Phase dient letztlich der Vorbereitung der Lese- und Rechtschreibentwicklung. Der eigentliche Schriftspracherwerb beginnt nach K. B. Günther (1986) jedoch erst mit dem Strategiewechsel in der logographemischen Phase.

Die zweite von K. B. Günther (1986) hinzugefügte Phase ist die integrativ-automatisierte Phase. Sie schließt sich an die orthographische Phase an und beschreibt den Automatisierungsprozess des Lesens und Rechtschreibens. „Sie stellt eigentlich keine neue Strategie mehr dar, sondern bezeichnet den schriftlichen Sprachgebrauch des kompetenten Lesers und Schreibers in einem autonomen und funktionsspezifischen Repräsentationssystem der Sprache“ (K. B. Günther, 1986, S. 43).

Die Entwicklungsmodelle von Frith (1985, 1986) und K. B. Günther (1986) beschreiben die qualitativen Veränderungen und verschiedenen Strategien in der kindlichen Entwicklung der Schriftsprache, wenn auch nur auf Wortebene. Obwohl sie empirisch nur unzureichend bestätigt sind, sind sie gleichwohl für die Schriftspracherwerbsforschung und Pädagogik nützlich (Scheerer-Neumann, 1996). Einige ihrer Leistungen sollen hier genannt sein:

1. Stufenmodelle stützen die Annahme, dass der Erwerb des Lesens und Schreibens ein integrativer Prozess ist und Unterricht den gleichzeitigen Erwerb fördern sollte (Dürscheid, 2006).

2. Stufenmodelle betonen, dass Kinder nicht erst mit dem Schuleintritt zu Lesern und Schreibern werden, sondern bereits unterschiedliche Vorerfahrungen gesammelt haben. Der Anfangsunterricht sollte dieser Heterogenität gerecht werden (Scheerer-Neumann, 1996).
3. Stufenmodelle gehen davon aus, dass der Erwerb des Lesens und Schreibens ein aktiver Lern- und Entwicklungsprozess ist, bei dem sich Kinder über unterschiedliche Strategien sukzessiv der Schriftsprache annähern. In diesem Sinne sind Fehler notwendige Entwicklungsmerkmale. Sie zeigen, wie ein Kind über Schriftsprache denkt und geben Hinweise auf den individuellen Entwicklungsstand (Brügelmann, 1986; Scheerer-Neumann, 1996).
4. Schließlich erlauben Stufenmodelle eine Vorhersage über die zukünftige Entwicklung. Der Unterricht sollte sich den unterschiedlichen Strategien, Entwicklungstempi und Lernständen der Kinder individuell anpassen (Scheerer-Neumann, 1996).

Andererseits wird an den Stufenmodellen von Frith (1985, 1986) und K. B. Günther (1986) auch Kritik geübt. Zum einen bestehen Zweifel an der logographemischen Phase, zum anderen sind die sequentielle Abfolge der Strategien und die Vorstellung von Frith, Strategien früherer Phasen stehen in höheren Phasen nicht mehr zur Verfügung, kritisch zu hinterfragen (Eichler, 1986). Die logographemische Phase wurde für den deutschsprachigen Raum schon oft in Frage gestellt (Eichler, 1986; Jansen, Mannhaupt & H. Marx, 1993; Nickel, 2006). Es gilt als gesichert, dass Stufenmodelle des Englischen nur begrenzt auf das Deutsche übertragen werden können. Es wird angenommen, dass beim Erwerb der deutschen Schriftsprache der Zugriff auf alphabetische Strategien schneller möglich ist als im Englischen, weil die deutsche Schriftsprache durch eine regelmäßige Zuordnung von Phonemen und Graphemen gekennzeichnet ist als das Englische (Jansen et al., 1993; Nickel, 2006). Außerdem muss berücksichtigt werden, dass der Unterricht Einfluss auf den Schriftspracherwerb ausübt. Der in Deutschland vorrangig angewendete analytisch-synthetische Leseunterricht beruht auf der alphabetischen Strategie. Daher werden logographemische Strategien sehr schnell überwunden (Scheerer-Neumann, 1996). Insofern nimmt die logographemische Strategie im deutschen Sprachraum nur einen sehr kleinen Teil ein oder ist gar nicht zu beobachten (Eichler, 1986; Jansen et al., 1993; Nickel, 2006). Eichler (1986) kritisiert weiterhin die von Frith (1985, 1986) postulierte Staffelung der Strategien und die Annahme, Strategien früherer Phasen stehen in höheren Phasen nicht mehr zur Verfügung. Auch wenn sich für das frühe Schreiben alphabetische und erst im späteren Entwicklungsverlauf orthographische Strategien feststellen lassen, geht Eichler

von einem parallelen Erwerb der Strategien aus. „Ich [Eichler] würde . . . eher für eine hierarchische Parallelität plädieren: man kann auf frühere Strategien zurück, und man kann Strategien parallel nutzen, aber die Reihenfolge des Erwerbs ist aufeinander aufbauend“ (Eichler, 1986, S. 239). Auch G. Thomé (2006) befürwortet die hierarchische Parallelität der Entwicklungsphasen, denn „kein Schülertext weist ausschließlich Schreibungen auf, die nur einer einzigen Phase zuzuordnen wären“ (S. 375). „In conclusion, it is clear that children are able to vary their strategy choice depending on the difficulty of the word they are attempting to spell . . .“ (Farrington-Flint, Stash & Stiller, 2008, S. 147).

3.3.2 Das Entwicklungsmodell des Rechtschreibkönnens von May (2002)

Das Entwicklungsmodell des Rechtschreibkönnens von May (2002) ist innerhalb des Diagnosekonzepts der Hamburger Schreibprobe (HSP) (May, 2002) entwickelt und empirisch bestätigt worden. Es stützt sich auf das Sechsstufenmodell von Frith (1985, 1986) und auf das Modell der Fertigkeitsentwicklung von H. L. Dreyfus und S. E. Dreyfus (1991).

Das Entwicklungsmodell geht davon aus, dass das Rechtschreiben auf unterschiedlichen *Strategien* basiert (May, 2001). May (2002) versteht unter Strategien individuelle Theorien und Vorstellungen über Schrift und Schreibweisen, die Einblicke in die kindliche Rechtschreibentwicklung geben und individuelle Entwicklungsverläufe erkennen lassen. In Anlehnung an Friths Modell (1985, 1986) werden diesem Entwicklungsmodell der logographemischen, alphabetischen und orthographischen Strategie die *morphematische Strategie* für das Schreiben von Sätzen und Texten sowie die *wortübergreifende Strategie* hinzugefügt. In der Entwicklung des Rechtschreibens erwerben Kinder die verschiedenen Strategien, wobei vorherige Strategien nicht aufgegeben, sondern in neuen Strategien weitergeführt und optimiert werden. Mit fortschreitender Rechtschreibentwicklung werden die einzelnen Strategien miteinander verbunden und bilden eine komplexe Gesamtstrategie. Es erfolgt eine zunehmende *Integration* einzelner Strategien. Damit greifen die verschiedenen Strategien beim Rechtschreiben ineinander, stützen sich gegenseitig und werden aufgabenspezifisch verwendet (May, 2002). Des Weiteren geht das Entwicklungsmodell davon aus, dass das Rechtschreiben auf *Regeln* basiert (May, 2001). Kinder erwerben Regeln über Schreibungen einzelner Wörter zumeist intuitiv durch das

Operieren mit Schriftsprache in Rechtschreibsituationen. Regeln helfen, orthographisch richtig zu schreiben. Sie entwickeln sich zu Strategien, wenn das Kind mehrere solcher Wortbildungsregeln entdeckt, verinnerlicht und routinemäßig danach handelt (May, 2002).

In Anlehnung an das von H. L. Dreyfus und S. E. Dreyfus (1991) entwickelte Modell der Fertigkeitsentwicklung vom Anfänger zum Experten, „vom regelgeleiteten ‚Know-that‘ zum erfahrungsbasierten Know-how“ (S. 41), beschreibt May (2002) die Rechtschreibentwicklung als einen stufenweisen Erwerb komplexer werdender Zugriffsweisen auf die Schriftsprache. Know-that umfasst Wissen über anzuwendende Regeln und Know-how umfasst Wissen, das durch Übung und Erfahrung erworben wurde (H. L. Dreyfus & S. E. Dreyfus, 1991). Entsprechend der fünf aufeinander aufbauenden Stufen der Fertigkeitsentwicklung vom Know-that zum Know-how durchlaufen nach May (2002) auch Kinder bei der Rechtschreibentwicklung die folgenden fünf Stufen: Anfänger (Neuling), fortgeschrittener Anfänger, Kompetenz, Gewandtheit und Expertentum. Die Stufen werden ebenso beim Erwerb der einzelnen Rechtschreibstrategien durchlaufen. Ist demnach ein bestimmtes Fertigniveau einer Strategie erreicht, kann auf bereits vorhandenen Strategien eine neue Strategie ausgebildet und differenziert werden (May, 2002). Das Entwicklungsmodell des Rechtschreibkönnens ist in Abbildung 8 schematisch zusammengefasst.

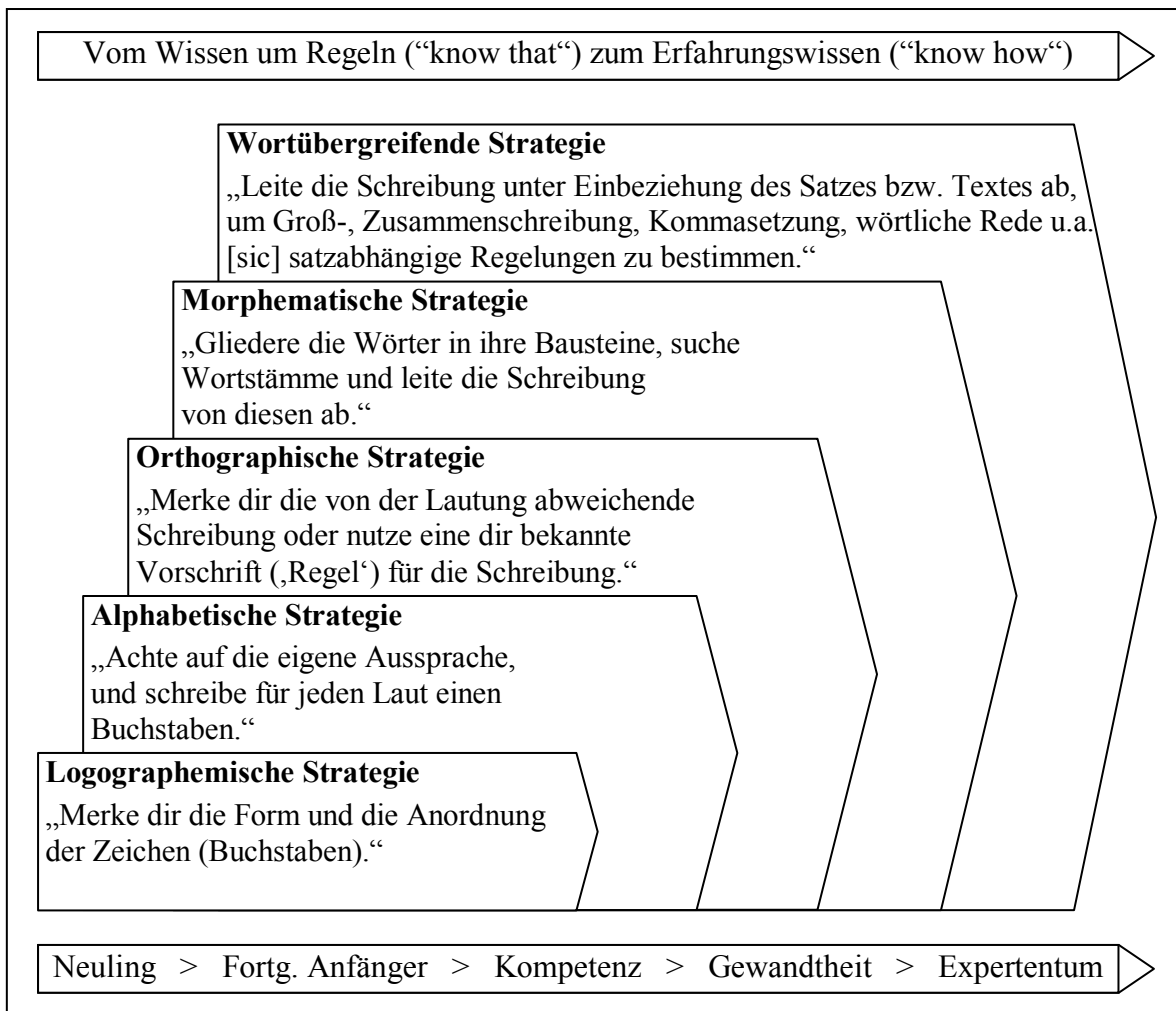


Abbildung 8: Entwicklung des Rechtschreibkönnens – Integration der grundlegenden Rechtschreibstrategien (May, 2002, S. 148)

Anmerkungen: Fortg. = Fortgeschrittener

Auf der Stufe des Neulings stützt sich das Kind beim Schreiben auf einzelne und isolierte Zugriffsweisen. Es kennt nur sehr allgemeine Handlungsregeln und wendet die logographemische und alphabetische Strategie an. Die logographemische Strategie beschreibt die Fähigkeit, sich Wörter und kurze Sätze unabhängig von der phonologischen Repräsentation zu merken und aufzuschreiben (May, 2001, 2002). Dabei befolgt das Kind die Selbstinstruktion „Merke dir die Form und die Anordnung der Zeichen (Buchstaben)“ (May, 2002, S. 27). Die alphabetische Strategie beschreibt die Fähigkeit, Phonemen Grapheme zuzuordnen und lautgetreu zu schreiben (May, 2001, 2002). Dabei befolgt das Kind die Selbstinstruktion „Achte auf die eigene Aussprache und schreibe für jeden Laut einen Buchstaben“ (May, 2002, S. 29). Innerhalb der Stufe des Neulings räumt May (2002) der logographemischen Strategie im Gegensatz zu Eichler (1986), Jansen et al. (1993) und Nickel (2006) eine wichtige Rolle in der Rechtschreibentwicklung ein. Sie wird nach May (2002) sowohl im beginnenden als auch im fortgeschrittenen

Rechtschreiberwerb für einige hundert bis tausend Wörter genutzt. Dieser Ansicht wird sich hier nicht angeschlossen. Es besteht kein Zweifel daran, dass für das Rechtschreibkönnen die Speicherung von Wörtern und Buchstabenkombinationen notwendig ist. Dazu gehören z. B. individuell bedeutsame Wörter, Flexionsmorpheme und Satzfunktionswörter wie Artikel und Präpositionen (Augst & Dehn, 2007). Dass diese Speicherung aber logographemisch, das heißt ohne Lautbezug und über Wortbilder, erfolgt, wird bezweifelt. Dies gilt auch dann noch, wenn May (2002) das logographemische Schreiben von der Wortbildtheorie abzugrenzen versucht. Es ist vielmehr, wie Augst und Dehn (2007), davon auszugehen, dass bei kompetenten Rechtschreibern *Schreibschemata* gespeichert sind, denn in der Rechtschreibentwicklung werden aus visuellen Wortbildern Schreibschemata, die aus bedeutungsunterscheidenden und von links nach rechts linear geordneten Graphemfolgen bestehen.

Auf der Stufe des fortgeschrittenen Anfängers hat das Kind schon einige Erfahrungen im Umgang Schriftsprache sammeln können. Es kann sich deshalb beim Rechtschreiben auf erweiterte und verfeinerte Zugriffsweisen stützen. Die alphabetische Strategie hat sich vervollkommnet und das Kind kennt alle ein- und mehrgliedrigen Grapheme. Die Strategie wird situationsbezogen angewendet und Wörter werden vollständig alphabetisch verschriftet. Da das Kind auf dieser Stufe um die unregelmäßigen Phonem-Graphem-Korrespondenzen weiß, versucht es orthographische Regeln für einzelne Wörter zu bilden. Die orthographischen Regeln überträgt es auch auf andere Wörter, wodurch es zu orthographischen und morphematischen Fehlern kommen kann. Diese Fehler werden vom Kind jedoch nicht bemerkt und korrigiert, da der beginnenden orthographischen und morphematischen Strategie noch bestimmte Fertigkeitsebenen fehlen (May, 2002).

Auf der Stufe der Kompetenz kennt das Kind alle wortbezogenen Strategien. Aufgrund von Erfahrungen kann das Kind auf diese vervollkommenen Strategien flexibel zurückgreifen, sie bewusst auswählen und anwenden. Zum Teil leitet das Kind beim Rechtschreiben schon wortübergreifende Aspekte her. Die wortübergreifende Strategie hat aber noch nicht das notwendige Fertigkeitsebene zur Integration erreicht. Deshalb stützt sich das Kind neben der alphabetischen auf die zwei weiteren wortbezogenen Strategien; die orthographische und die morphematische Strategie. Die orthographische Strategie beschreibt die Fähigkeit, einfache Phonem-Graphem-Beziehungen entsprechend orthographischer Regelungen zu modifizieren. Es sind zwei orthographische Regelungen

zu unterscheiden; die *Merkelemente* und die *Regelemente*. Bei Merkelementen muss sich die von der Artikulation abweichende Schreibung gemerkt werden. Bei Regelementen kann die Schreibung aus zugrunde liegenden Regeln, z. B. Pluralbildung bei <Hand>, abgeleitet werden (May, 2001, 2002). Bei der orthographischen Strategie befolgt das Kind die Selbstinstruktion „Merke dir die von der Lautung abweichende Schreibung der Wortstelle und/oder nutze eine dir bekannte Vorschrift (‘Regel’) für die Schreibung“ (May, 2002, S. 30). Die morphematische Strategie beschreibt die Fähigkeit, Wörter in Morpheme zu gliedern und dadurch Schreibungen abzuleiten. Das Kind kennt mit der morphematischen Strategie das Prinzip der Morphemkonstanz und kann dadurch den Wortbestandteilen orthographische Regeln zuordnen. Es kann „die Wortschreibungen sozusagen nach einem ‚Baukastenprinzip‘ rekonstruieren, der kognitive Aufwand beim Schreiben wird ökonomisch genutzt und das Gedächtnis wird dadurch enorm entlastet“ (May, 2002, S. 31). Beim morphematischen Schreiben werden die Wörter hinsichtlich ihrer *morphosemantischen* und *morphologischen* Wortbildung analysiert. Die morphosemantische Struktur fragt nach der Bedeutung eines Wortes und seiner Grundmorpheme und kann dadurch die Schreibung ableiten, z. B. *Staub* und *staubig*. Die morphologische Struktur fragt nach den Morphemen eines Wortes und kann dadurch die Schreibung ableiten, z. B. *Verkäuferin* und *Ver-käuf-er-in* (May, 2001, 2002). Bei der morphematischen Strategie befolgt das Kind die Selbstinstruktion „Gliedere die Wörter in ihre Bausteine, suche nach verwandten Wortstämmen und leite die Schreibung von diesen ab“ (May, 2002, S. 32). Die morphematische Strategie ist die umfangreichste und wichtigste Strategie. „Schreiber, die dieses Wissen haben, können im Zweifelsfall orthographisch relevante Verfahrensweisen anwenden, um die Schreibweise eines Wortes richtig zu entscheiden“ (May, 2002, S. 32).

Auf der Stufe der Gewandtheit verfügt das Kind über vielfältige Erfahrungen beim Rechtschreiben. Die wortübergreifende Strategie hat sich vervollkommen und greift mit den anderen Strategien ineinander. Sie kann durch das Kind flexibel angewendet werden. Die wortübergreifende Strategie beschreibt die Fähigkeit, beim Schreiben weitere linguistische Aspekte zu beachten. Dazu gehören u. a. die Wortart für die Groß- und Kleinschreibung, die Wortsemantik für die Zusammen- und Getrennschreibung, die Satzgrammatik für die Kommaschreibung und die grammatische Kongruenz für die Deklination (May, 2001, 2002). Bei der wortübergreifenden Strategie befolgt das Kind die Selbstinstruktion „Leite die Schreibung unter Einbeziehung des Satzes bzw. Textes ab,

um Groß-, Zusammenschreibung, Kommasetzung, wörtliche Rede u.a. [*sic*] satzabhängige Regelungen zu bestimmen ““(May, 2002, S. 148).

Auf der Stufe des Expertentums hat das Kind einen reichen Erfahrungsschatz von Schriftsprache und Rechtschreiben, sodass es zumeist intuitiv handelt. „Der ‚Rechtschreibexperte‘ schreibt ‚mit dem Gefühl für das Richtige‘ und überlegt nicht, warum ein Wort so geschrieben oder ein Komma an dieser Stelle im Satz gesetzt wird“ (May, 2002, S. 147).

Das Entwicklungsmodell des Rechtschreibkönnens von May (2002) hält an den wortbezogenen Strategien von Frith (1985, 1986) fest und erweitert diese um die morphematische und wortübergreifende Strategie. Die morphematische Strategie ist neben der alphabetischen Strategie die wichtigste Rechtschreibstrategie, da sie ökonomisch ist und nicht einzelwortbezogen gilt. Das Modell betont jedoch nicht die Vielzahl orthographischer Regeln und Strategien, die sich ein Kind in der Rechtschreibentwicklung aneignet. Der Fokus des Modells liegt vielmehr auf der Vielzahl unterschiedlicher Erfahrungen, die ein Kind sammeln sollte. Das bedeutet, dass für die Herausbildung und Integration der Rechtschreibstrategien viel geschrieben (und gelesen) werden muss. Allein dadurch schreitet das Regelwissen, das Know-that, zum Erfahrungswissen, Know-how, voran und ermöglicht ein situationsspezifisches und flexibles rechtschreibliches Handeln (May, 2002).

3.3.3 Das Stufenmodell der Rechtschreibentwicklung von Scheerer-Neumann (2008)

Aus der Kritik an „harte[n]‘ Versionen von Stufenmodellen“ (Scheerer-Neumann, 1996, S. 1154), die eine irreversible Stufenabfolge annehmen, ist das Entwicklungsmodell von Scheerer-Neumann (2008) entstanden. Es ist ein Versuch, die dominierenden Strategien in der Rechtschreibentwicklung zu beschreiben und dabei Raum für individuelle Entwicklungen und aufgabenspezifische Varianten zu lassen (Scheerer-Neumann, 1996). Das Stufenmodell der Rechtschreibentwicklung von Scheerer-Neumann (2008) beschreibt *Entwicklungsstufen*. Scheerer-Neumann (2007) betont, diesen Begriff nicht im engeren Sinn zu verstehen, sondern als theoretisches Konstrukt, welches den Entwicklungsverlauf

strukturiert. In diesem Sinn beschreiben Entwicklungsstufen Phasen, in denen bestimmte Rechtschreibstrategien dominieren und später von anderen dominierenden Strategien abgelöst werden. In den Phasen selbst finden auch Veränderungen statt, die den Übergang zur nächsthöheren Phase vorbereiten. Das Stufenmodell von Scheerer-Neumann (2008) beruht auf ihrem Prozessmodell des Rechtschreibens (Scheerer-Neumann, 2003), das annimmt, Rechtschreiben erfolgt über Einsichten in orthographische Strukturen und über den Erwerb wortspezifischer Informationen, sogenannter Lernwörter. Beide Gedächtnisspeicher, das Regelwissen, das sich in den dominierenden Rechtschreibstrategien zeigt, und die Lernwörter, sind innerhalb des Stufenmodells miteinander verbunden. Demzufolge macht das Stufenmodell neben dem Entwicklungsverlauf der Strategien auch deutlich, welchen Einfluss Rechtschreibstrategien auf den Erwerb von Lernwörtern haben (Scheerer-Neumann, 2006). Tabelle 2 zeigt das Stufenmodell der Rechtschreibentwicklung von Scheerer-Neumann (2008).

Tabelle 2: Entwicklungsstufen beim Erwerb der Rechtschreibung (Scheerer-Neumann, 2008, S. 166)

Entwicklungsstufen/Rechtschreibstrategien (Beispiele für Schreibungen)	Lernwörter
Logographemisch (beliebige Buchstabenfolgen, Kritzeln)	Eigener Name und andere Namen, nur sehr wenige Lernwörter möglich
Beginnend alphabetisch (MT für Mutter)	Immer noch sehr wenige Lernwörter
Entfaltet alphabetisch (Woke für Wolke, Bume für Blume)	Erwerb von Lernwörtern schon etwas leichter
Voll entfaltete alphabetische Strategie (Blume, Kint, Fel, Sane)	Lernwörter können jetzt recht gut erworben werden
Beginnende Einsicht in orthographische Strukturen	Lernwörter mit den entsprechenden orthographischen Strukturen können leicht erworben werden
Zunehmend weitere Einsichten in orthographische Strukturen	Erwerb von Lernwörtern wird weiter erleichtert
Überwiegen des Abrufens von Lernwörtern über Konstruktionen	

Vor dem eigentlichen Schriftspracherwerb steht die von K. B. Günther (1986) beschriebene präliteral-symbolische Stufe. In dieser Phase vor dem Schuleintritt erfahren Kinder, dass Buchstaben besondere Zeichen sind und sich von Bildern unterscheiden. Mit Kritzeleien, die zum Teil schon Buchstabenfragmenten ähneln, ahmen sie das Schreiben nach. Kritzeleien spiegeln die ersten kindlichen Einsichten in die Schrift wider, denn in der

Regel sind sie linear ausgerichtet und wie die Schreibrichtungen miteinander verbunden (Scheerer-Neumann, 2007, 2008).

Mit der *logographemischen Strategie* als erste Entwicklungsstufe des Rechtschreiberwerbs haben Kinder die kommunikative Funktion der Schrift erkannt. Zwar ist das Kritzeln bei spontanen Schreibversuchen noch vorherrschend, aber durch Abschreiben und Auswendiglernen können bereits ganze Wörter wie der eigene Name oder <MAMA> und <PAPA> geschrieben werden. Kinder können außerdem schon einige Buchstabennamen nennen. Die Zuordnung von Phonemen und Graphemen erfolgt aber nicht. Das logographemische Vorgehen stellt hohe Anforderungen an das Gedächtnis, da Wörter nicht an ihrer Lautstruktur erkannt, sondern nur als Lernwörter erworben werden können. Aufgrund dessen kommt es beim Schreiben häufig zu Auslassungen und Umstellungen von Graphemen (Scheerer-Neumann, 2007, 2008).

Die sich anschließende *alphabetische Strategie* ist der größte Schritt in der Rechtschreibentwicklung, weil Kinder den Zusammenhang von gesprochener und geschriebener Sprache erfassen. Sie erkennen zumeist im frühen ersten Schuljahr das Prinzip der Phonem-Graphem-Korrespondenz. Die alphabetische Rechtschreibstrategie erstreckt sich über drei Entwicklungsstufen (Scheerer-Neumann, 2007, 2008).

Mit der *beginnenden alphabetischen Strategie* können Kinder erstmals Wörter ohne Vorlage konstruieren. Bei den ersten Verschriftlichungen kommt es noch zu vielen Graphemauslassungen innerhalb des Wortes. Sie verschriften fast immer das Anfangsphonem, lassen aber häufig Vokale aus, z. B. <MT> für *Mutter*, <ELF> für *Elefant* (Scheerer-Neumann, 2007, 2008). Diese ersten alphabetischen Schreibungen werden „*konsonantische Skelettschreibungen*“ (Scheerer-Neumann, 2007, S. 546) genannt. Auf der Stufe der beginnenden alphabetischen Strategie fällt der Erwerb von Lernwörtern etwas leichter, da Kinder erste Einsichten in die Phonem-Graphem-Korrespondenzen gewinnen. Gleichzeitig stellen aber das alphabetische Vorgehen und die Phonemanalyse, die sich nur entwickeln können, wenn die phonologische Bewusstheit gut ausgebildet ist, für die Kinder hohe Anforderungen dar (Scheerer-Neumann, 2006, 2007, 2008).

Während des 1. Schuljahres vervollständigen sich die kindlichen Schreibungen, sodass nur noch wenige Grapheme ausgelassen werden. Fast jedem Phonem wird ein Graphem

zugeordnet. Nur noch bei Konsonantenhäufungen und in langen Wörtern werden Grapheme ausgelassen, z. B. <KOKODIL> für *Krokodil*. Damit erreichen Kinder die *entfaltete alphabetische Strategie* (Scheerer-Neumann, 2007, 2008). Zu diesem Zeitpunkt begleiten Kinder das Schreiben durch mehrmaliges Sprechen der Wörter, um ihre Aufmerksamkeit vollständig auf die Phoneme zu richten. Das Einprägen von Lernwörtern gelingt hier bereits einfacher als zuvor, weil es mehr phonemisch unterstützt wird als bei der beginnenden alphabetischen Strategie (Scheerer-Neumann, 2007).

Die *voll entfaltete alphabetische Strategie* erreichen Kinder mit verstärkter Übung der Phonemanalyse im 2. Schuljahr. Die Phonemanalyse und die Phonem-Graphem-Korrespondenz gelingen den Kindern mit zunehmender Geläufigkeit. Dadurch können Lernwörter recht gut erworben werden, denn nur noch wortspezifische Merkmale wie z. B. das *ß* in *Fuß* oder Abweichungen von der Phonem-Graphem-Korrespondenz müssen sich gemerkt werden. Allerdings treten in dieser Stufe gehäuft Rechtschreibfehler auf, die durch die dominierende alphabetische Strategie hervorgerufen werden. So werden bekannte Lernwörter alphabetisch verschriftet, obwohl Kinder ihre spezifischen Schwierigkeiten bereits kennen, z. B. <komt> für *kommt* (Scheerer-Neumann, 2007, 2008).

Mit der *orthographischen Strategie*, die sich nicht unmittelbar an die entfaltete alphabetische Strategie anschließen muss, sondern bereits parallel dazu auftreten kann, gewinnen Kinder Einblicke in orthographische Regelmäßigkeiten. Sie erlernen u. a. die Groß- und Kleinschreibung, das Prinzip der Morphemkonstanz und erwerben die Endmorpheme *-en* und *-er*. Die Vielzahl orthographischer Regelmäßigkeiten und Strukturen erwerben Kinder über mehrere Jahre. Deshalb wird die orthographische Strategie in *beginnende* und *zunehmende Einsichten in orthographische Strukturen* unterteilt (Scheerer-Neumann, 2007, 2008). Auf der Stufe der beginnenden orthographischen Strategie wird die alphabetische Strategie mit wortspezifischen Informationen kombiniert. Alphabetische Schreibungen werden so durch bereits bekannte orthographische Elemente korrigiert. Mit der fortschreitenden Rechtschreibentwicklung gewinnen Kinder immer neue Einsichten in orthographische Strukturen, vertiefen diese und werden für besondere Schreibungen sensibilisiert (Scheerer-Neumann, 1987). „Ganz allgemein erleichtern strukturelle Regelmäßigkeiten den Lernprozeß [*sic*]; sie reduzieren die Information, die ‚ganz neu‘ eingepägt werden muß [*sic*]“ (Scheerer-Neumann, 1987, S. 212). Somit wird der Erwerb von Lernwörtern weiter erleichtert (Scheerer-Neumann,

2008).

Mit dem *Überwiegen des Abrufens von Lernwörtern* als letzte Entwicklungsstufe der Rechtschreibentwicklung verändert sich das Verhältnis zwischen dem Konstruieren von Wörtern und dem Abrufen von Lernwörtern zugunsten des Abrufprozesses, welcher weitgehend automatisiert wird. Allerdings bleibt die alphabetische Strategie auch bei Erwachsenen zur Handlungssteuerung und für das innere Mitsprechen beim Schreiben weiterhin vorhanden (Scheerer-Neumann, 2007, 2008).

Das Stufenmodell der Rechtschreibentwicklung von Scheerer-Neumann (2008) hält an den Grundzügen der Entwicklungsmodelle von Frith (1985, 1986) und K. B. Günther (1986) fest. Scheerer-Neumann (2007, 2008) legt aber den Schwerpunkt auf die alphabetische Stufe und unterteilt diese, weil die Entwicklung der alphabetischen Strategie komplex ist. Das Stufenmodell betont zudem, dass die alphabetische Strategie selbst während der orthographischen Entwicklungsstufen nicht aufgegeben wird. Das bedeutet, dass die orthographische Strategie parallel zur alphabetischen Strategie auftreten kann und immer ein Zugriff auf die alphabetische Strategie möglich ist. Das Stufenmodell zeigt außerdem, inwieweit der Erwerb der Lernwörter von der jeweiligen Entwicklungsstufe abhängt. Je weiter die Rechtschreibentwicklung voranschreitet, desto leichter können Kinder alphabetisch schreiben und sich Lernwörter merken. „Lernwörter werden dann leichter gelernt und weniger schnell vergessen, wenn sie an alphabetischen oder orthografischen Strukturen verankert sind. Deshalb ist es so wichtig, den Stand der Entwicklung der Schreibstrategien zu kennen“ (Scheerer-Neumann, Schnitzler & Ritter, 2009).

3.3.4 Das Kompetenzentwicklungsmodell des Rechtschreibens von Löffler und Meyer-Schepers (2008)

Das Kompetenzentwicklungsmodell des Rechtschreibens von Löffler und Meyer-Schepers (2008) ist im Rahmen der erweiterten Internationalen Grundschul-Lese-Untersuchung (IGLU-E) entwickelt worden (Löffler & Meyer-Schepers, 2005, 2006). Das Kompetenzentwicklungsmodell ist ein *2-Fähigkeiten-* und *2-Ebenenmodell* (2-2-Modell) der Rechtschreibkompetenz und gilt für die Klassenstufen eins bis vier bzw. sechs. Es differenziert die allgemeine Rechtschreibkompetenz in zwei grundlegende Dimensionen,

in die *phonographische* (lautanalytische) und *grammatische Fähigkeit*. Zudem werden zwei Ausprägungsebenen für beide Fähigkeiten unterschieden, eine *elementare* und eine *erweiterte Niveaustufe*. Im Gegensatz zu den bisher aufgezeigten Entwicklungsmodellen geht das Kompetenzentwicklungsmodell nicht von einem stufenweisen Erwerb dominierender Rechtschreibstrategien aus. Das Modell beschreibt die Rechtschreibentwicklung als parallele Entwicklung der phonographischen und grammatischen Fähigkeit. Beide Fähigkeiten entwickeln sich, ähnlich dem Modell von May (2002), von der elementaren zur erweiterten Niveaustufe und verzahnen sich zur Gesamtkompetenz Rechtschreibfähigkeit. Je nach Ausprägungsgrad und Niveaustufe der Fähigkeiten können Kinder Rechtschreibaufgaben mit unterschiedlich hohem Anforderungsniveau bewältigen. Die elementaren lautanalytischen und grammatischen Fähigkeiten sollten in den Klassenstufen eins und zwei, die erweiterten Fähigkeiten in den Klassenstufen drei bis vier bzw. sechs erworben werden (Löffler & Meyer-Schepers, 2008; Löffler, Meyer-Schepers & Stubbe, 2008; Voss, Löffler, Meyer-Schepers, Meckel & Kowalski, 2008). Das Kompetenzentwicklungsmodell des Rechtschreibens ist in Abbildung 9 dargestellt.

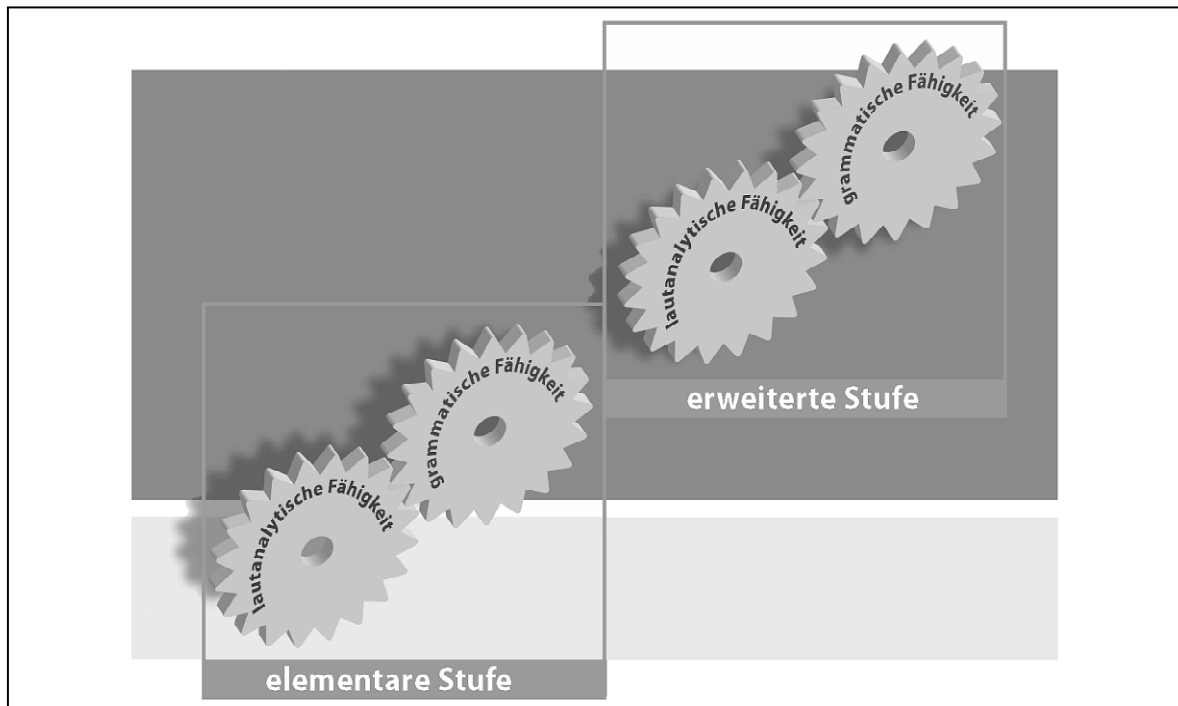


Abbildung 9: Kompetenzentwicklungsmodell des Rechtschreibens (in Anlehnung an Lischeid, 2007)

Die lautanalytische Fähigkeit erfordert grundlegende Einsichten in die Funktion und den Aufbau von Schrift und Sprache sowie in das phonematische Prinzip. Mit der

lautanalytischen Fähigkeit können Kinder den Lautbestand der gesprochenen Sprache hinsichtlich der Phoneme analysieren. Darüber hinaus erkennen sie, dass Phoneme durch Grapheme repräsentiert werden. Für die lautanalytische Fähigkeit wird eine Niveaueinteilung in eine elementare und erweiterte Stufe vorgenommen. Auf der elementaren Stufe beherrschen Kinder einfache Phonem-Graphem-Korrespondenzen und bestimmte Abweichungen. Sie können einem Phonem ein Graphem oder ein mehrgliedriges Graphem, z. B. /p/ und ⟨p⟩ oder /ʃ/ und ⟨sch⟩, zuordnen. Umgekehrt können Sie einem mehrgliedrigen Phonem ein Graphem, z. B. /ts/ und ⟨z⟩, zuordnen. Zudem haben sie auf der elementaren Stufe erste Einsichten gewonnen, dass einem Phonem mehrere Grapheme zugeordnet werden können. Auf der erweiterten Stufe beherrschen Kinder die graphematische Repräsentation der Vokalquantität durch Kürze- und Dehnungszeichen. Sie haben damit die bedeutungsunterscheidende Opposition von Länge und Kürze eines Vokals erfasst. Sie können die Kürze eines Vokals durch Dopplungs- und Kürzezeichen wiedergeben, z. B. ['kasə] und *Kasse* sowie ['bakən] und *backen*. Sie können die Länge eines Vokals durch Dehnungszeichen wiedergeben, z. B. [ba:n] und *Bahn*, [mo:r] und *Moor* sowie [ʃpi:l] und *Spiel* (Löffler & Meyer-Schepers, 2005, 2008; Löffler et al., 2008; Voss et al., 2008).

Die grammatische Fähigkeit umfasst Fähigkeiten, die für Wort- und Satzschreibungen notwendig und nicht auf der Grundlage des Lautprinzips zu erschließen sind. Die grammatische Fähigkeit korrigiert und ergänzt die lautanalytischen Fähigkeiten. Sie erfordert Einsichten in die übrigen Bereiche der phonologischen und semantischen Grundprinzipien. Für die grammatische Fähigkeit wird eine Niveaueinteilung in eine elementare und erweiterte Stufe vorgenommen. Auf der elementaren Stufe beherrschen Kinder die Schreibung von Vor- und Endsilben, z. B. *be-*, *auf-*, *an-*, *-en*, *-el* und *-er*. Im Bereich der Morphemkonstanz haben sie Einsichten in elementare Ableitungen, z. B. *Ast* und *Äste* sowie *Baum* und *Bäume*, und Auslautverhärtungen der Verschlusslaute gewonnen. Im Bereich der Groß- und Kleinschreibung beherrschen sie konkrete, das heißt seh- und greifbare, Nomen und können Wortgrenzen einhalten. Zudem gelingt ihnen die Großschreibung am Satzanfang. Auf der erweiterten Stufe beherrschen Kinder im Bereich der Morphemkonstanz erweiterte Ableitungen, z. B. das *ß* nach langem Vokal wie in *Maß* oder das silbenanlautende *h* wie in *gehen*. Im Bereich der Großschreibung beherrschen sie nun auch abstrakte, das heißt vorstellbare, Nomen wie *Pause* und Nominalisierungen wie *das Gute* (Löffler & Meyer-Schepers, 2005, 2008; Löffler et al., 2008; Voss et al., 2008).

Voss et al. (2008) liefern in einer empirischen Untersuchung Belege für die empirische Evidenz des Kompetenzentwicklungsmodells. Sie bestätigen zum einen die in dem Modell vorgenommene Unterscheidung zwischen lautanalytischer und grammatischer Kompetenz. Zum anderen bestätigen sie auch die Niveaueinteilung der Kompetenzen. Das 2-2-Modell ist ein „**integratives Schriftsprachmodell**“ (Löffler & Meyer-Schepers, 2007, S. 182). Es betrachtet Rechtschreibentwicklung als parallelen Erwerb, Verzahnung und Integration lautanalytischer und grammatischer Fähigkeiten. Beide Fähigkeiten sind von Beginn an für korrektes Rechtschreiben notwendig. Hilft die lautanalytische Fähigkeit für eine bestimmte Schreibung nicht weiter, wird auf die grammatische Fähigkeit der jeweiligen Stufe zurückgegriffen. Nur so können Kinder die orthographisch richtige Schreibung für z. B. die Lautung ['kʏstə] (<küsste> oder <Küste>) wählen (Löffler & Meyer-Schepers, 2005, 2007; Löffler, Meyer-Schepers & Lischeid, 2007). Das 2-2-Modell ist zugleich ein „*orthografiedidaktisches Modell*“ (Löffler et al., 2007, S. 269). Es betrachtet Rechtschreibung als einen zu unterrichtenden Lernbereich, da Rechtschreiberwerb keine Selbstverständlichkeit ist. Unterricht und sichernde Übungen sind für den Erwerb der Rechtschreibung notwendig. „Das Plädoyer des gutschrift|Kompetenzmodells [gutschrift ist ein Referenzzentrum für die Diagnostik der Schriftkompetenz unter der Leitung von Löffler und Meyer-Schepers] ist aber, dass die Vermittlung orthografiedidaktischen Wissens auf den integrativen Aufbau beider Fähigkeitsdimensionen abzielen muss, da von Anfang an neben die phonographische die grammatische Kompetenz für die korrekte Schreibung hinzutreten muss“ (Löffler et al., 2007, S. 270).

3.3.5 Zusammenfassung und Schlussfolgerungen

Der Überblick über die meist diskutierten Modelle der Rechtschreibentwicklung im deutschsprachigen Raum verdeutlicht, dass Rechtschreiberwerb ein Entwicklungsprozess ist, der sich in den unterschiedlichen Zugriffsweisen auf die Schriftsprache ausdrückt. Über die Abfolge der unterschiedlichen Zugriffsweisen auf die Schriftsprache sind sich die Modelle der Rechtschreibentwicklung einig. Es wird angenommen, dass die Rechtschreibentwicklung bereits vor dem Schuleintritt beginnt, wobei der logographemischen Phase im deutschsprachigen Raum deutlich weniger Bedeutung zukommt als im englischsprachigen Raum. Ausgehend vom lautorientierten Schreiben entwickelt sich über unterschiedliche Stufen, Phasen und Strategien die voll entfaltete

Rechtschreibfähigkeit, bei der die Prinzipien und Regeln der Rechtschreibung berücksichtigt werden. Konsens besteht auch darüber, dass frühere Rechtschreibstrategien in neu erworbenen und dominierenden Strategien fortgeführt werden. Somit können Kinder in Rechtschreibsituationen mit unterschiedlichen Anforderungsbereichen auf unterschiedliche Strategien zurückgreifen.

Entwicklungsmodelle beschreiben idealtypische Entwicklungsverläufe und machen durch die Annahme von Stufen, Phasen und Strategien die Rechtschreibentwicklung transparenter. Durch sie lässt sich der Entwicklungsstand eines Kindes bestimmen, aber nicht der individuelle Aneignungsweg, z. B. wie lange ein Kind eine bestimmte Strategie nutzen wird, voraussagen. Gerade die zeitliche Dauer einer Strategie ist individuell sehr verschieden und zudem abhängig von individuellen Lernvoraussetzungen und schulischen Lern- und Lehrprozessen. Deshalb ist es wichtig, dass Entwicklungsmodelle Raum für individuelle Entwicklungen lassen und dass immer zwischen idealtypischer und individueller Entwicklung unterschieden wird (Kirschhock, 2004; Nickel, 2006). Insofern dienen Entwicklungsmodelle nur bedingt der pädagogischen Orientierung und didaktischen Unterrichtsplanung. „Entwicklungsmodelle würden falsch verstanden, wenn sie als Grundlage eines didaktischen Konzepts eingesetzt würden und die Lernwege – dem idealtypisch angenommenen Lernprozess folgend – festlegen und damit einengen würden“ (Nickel, 2006, S. 102). Jedoch bilden sie einen guten Rahmen für Diagnose und Förderung. Sie ermöglichen die Einschätzung des Entwicklungsstandes und das Erkennen von Schwierigkeiten im Entwicklungsprozess, aber nicht das Erklären dieser (Nickel, 2006).

Aus den bisherigen Ausführungen ergibt sich für die Entwicklung der FE-RS 2 Folgendes: Für das dieser Arbeit zugrunde liegende Modell der Rechtschreibentwicklung soll die im Abschnitt 3.2.2 erläuterte Unterscheidung von Graphematik und Orthographie (Neef, 2005b) Berücksichtigung finden. Neef (2005a) nimmt an, dass sich Kinder beim Erwerb der Rechtschreibung in einem ersten Schritt graphematisches Können und in einem zweiten Schritt orthographisches Können aneignen. Diesem Postulat wird das Kompetenzentwicklungsmodell des Rechtschreibens von Löffler und Meyer-Schepers (2008) gerecht. Für das Ziel dieser Arbeit, ein sensitives Verfahren zu entwickeln, scheint es aber allein nicht differenziert genug. Aufgrund dessen soll es mit dem Stufenmodell der Rechtschreibentwicklung von Scheerer-Neumann (2008) und dem Entwicklungsmodell des Rechtschreibkönnens von May (2002) kombiniert werden. Das Entwicklungsmodell

von Scheerer-Neumann (2008) besticht durch differenzierte Aussagen zur alphabetischen Strategie, die als Rechtschreibstrategie die zweite Klassenstufe dominiert (Naumann, 2008; Scheerer-Neumann, 2003, 2006). Das Modell von May (2002) hingegen berücksichtigt differenziert die erweiterten Rechtschreibfähigkeiten, die zunehmend mit der alphabetischen Strategie integriert und verzahnt werden. Das dieser Arbeit zugrunde liegende Modell der Rechtschreibentwicklung kann somit als Synopse des Postulats von Neef (2005a, b) und den Modellen von Löffler und Meyer-Schepers (2008), May (2002) und Scheerer-Neumann (2008) verstanden werden. Diese theoretische Zusammenführung ist in Tabelle 3 wiedergegeben.

Tabelle 3: Entwicklungsmodell des graphematischen und orthographischen Rechtschreibkönnens (in Anlehnung an Löffler & Meyer-Schepers, 2008; May, 2002; Neef, 2005a, b; Scheerer-Neumann, 2006, 2008)

	Teilfähigkeit der Rechtschreibleistung	Dominierende Rechtschreibstrategie	Beherrschungsgrad in den Klassen					
			1	2	3	4	5	6
Elementare Stufe	Graphematisches Können	<i>Beginnend alphabetische Strategie</i> (erste, einfache PGK) <i>Entfaltet alphabetische Strategie</i> (fast vollständige PGK) <i>Voll entfaltete alphabetische Strategie</i> (vollständige PGK)						
	Orthographisches Können	<i>Beginnend orthographische Strategie</i> (Großschreibung konkreter Nomen und des Satzanfangs, elementare Ableitung, Vor-/Endsilben, Auslautverhärtung)						
Erweiterte Stufe	Graphematisches Können	<i>Zunehmend orthographische Strategie</i> (Vokalquantität)						
	Orthographisches Können	<i>Morphematische Strategie</i> (Einsichten in morphosemantische und morphologische Zusammenhänge) <i>Wortübergreifende Strategie</i> (Großschreibung abstrakter Nomen und von Nominalisierungen, das/dass, Kommasetzung)						

Anmerkungen: Die Graufärbung der Felder zeigt den Beherrschungsgrad der Strategien in den Klassenstufen an. Schwarz bedeutet, die meisten Schüler beherrschen die jeweilige Strategie, etc.. PGK = Phonem-Graphem-Korrespondenz

Vor dem Hintergrund des Entwicklungsmodells soll das zu entwickelnde Verfahren die Entwicklung der Rechtschreibleistungen der Schüler der 2. Klassenstufe dokumentieren. Es soll vorwiegend elementares graphematisches und orthographisches Können erfassen,

da dieses die zweite Klassenstufe dominiert und zu diesem Zeitpunkt von fast allen Schülern weitgehend beherrscht wird. Zudem soll es erste erweiterte graphematische Kenntnisse erfassen, da diese in der 2. Klassenstufe zunehmend mit dem elementaren Rechtschreibkönnen verbunden werden.

Obwohl sich mit Hilfe von Entwicklungsmodellen der Entwicklungsstand eines Kindes bestimmen lässt, bleibt zu berücksichtigen, dass von der Schreibung eines Wortes nicht direkt auf die kognitive Tätigkeit beim Rechtschreiben geschlossen werden kann. Schreibungen können nämlich unterschiedlich zustande kommen; durch Abschreiben, einfaches Merken oder Konstruieren (May, 2002). Welche kognitiven Teilprozesse am Rechtschreiben beteiligt sind, wird nachstehend erläutert. Der Abschnitt 3.4 gibt einen Überblick über Modelle des Rechtschreibprozesses.

3.4 Modelle des Rechtschreibprozesses

Modelle des Rechtschreibprozesses beschreiben die kognitiven Prozesse und das Zusammenwirken unterschiedlicher Elemente während des Rechtschreibens. Auch wenn Modelle des Rechtschreibprozesses denen des Lesens ähneln, bestehen hinsichtlich der Verarbeitungsprozesse und Anforderungen Unterschiede (Klicpera, Schabmann & Gasteiger-Klicpera, 2007). Das Rechtschreiben unterscheidet sich vom Lesen in folgenden Punkten:

1. Rechtschreiben ist ein langsamerer Prozess, bei dem Inhalte aus dem Gedächtnis abgerufen und behalten werden müssen, was aufwendiger und störungsanfälliger ist.
2. Es gibt mit dem freien Schreiben von Wörtern und Texten, dem Abschreiben, dem Schreiben nach Diktat sowie dem Korrigieren eigener und fremder Verschriftungen mehr Anforderungsbereiche als beim Lesen.
3. Die Phonem-Graphem-Korrespondenzen sind viel unregelmäßiger als umgekehrt, was bedeutet, dass das lautorientierte Lesen schneller und reibungsloser verläuft als das lautorientierte Rechtschreiben (Hinney et al., 2008; Klicpera et al., 2007; Mannhaupt, 2008; Nickel, 2006).

Rechtschreiben und Lesen sind folglich keine spiegelbildlichen Prozesse (Klicpera et al., 2007). Dennoch besteht ein enger Zusammenhang zwischen der Rechtschreib- und Leseleistung, denn gute Rechtschreiber haben bessere Leseleistungen als

rechtschreibschwache Schüler (Löffler et al., 2008). Die Längsschnittstudie von Klicpera, Schabmann und Gasteiger-Klicpera (2006) zeigt zudem, dass das Lesen in den ersten beiden Klassenstufen Einfluss auf das Rechtschreiben hat. Von der 2. zur 4. Klassenstufe kehrt sich die Interaktion zwischen den Leistungen um. Da es sich beim Lesen und Rechtschreiben um zwei verschiedene Prozesse handelt, auch wenn es Zusammenhänge zwischen beiden gibt, muss zwischen Modellen des Lese- und Rechtschreibprozesses unterschieden werden. Bereits 1973 wurde ein Rechtschreibprozessmodell von D. P. Simon und H. A. Simon entwickelt. Später wurden *Zwei-Wege-Modelle* und *Netzwerkmodelle*, auch als *konnektionistische Modelle* bezeichnet, entwickelt.

Als eines der zuerst entwickelten Modelle des Rechtschreibprozesses wird im Folgenden das Modell von D. P. Simon und H. A. Simon (1973) beschrieben. Als Beispiele für Zwei-Wege-Modelle des Rechtschreibprozesses werden die Modelle von Barry (1994) und Scheerer-Neumann (2003) dargelegt. Als Beispiele für Netzwerkmodelle des Rechtschreibprozesses werden die Modelle von Brown und Loosemore (1994) sowie Olson und Caramazza (1994) erläutert.

3.4.1 Das Generator-Test-Prozessmodell von D. P. Simon und H. A. Simon (1973)

Das Rechtschreibprozessmodell von D. P. Simon und H. A. Simon (1973) ist eines der älteren Modelle. Es findet in nahezu jeder Literatur über Rechtschreibmodelle Erwähnung und bildet die Grundlage für spätere Prozessmodelle.

D. P. Simon und H. A. Simon (1973) nehmen unterschiedliche Wege an, die beim Schreiben eines Wortes gegangen werden können. Der Rechtschreibprozess kann zum einen über das *Gedächtnis* erfolgen. Dabei wird sich über die Bedeutung und die Aussprache des Wortes an dessen Schreibung erinnert. Dies ist der einfachste Weg. Zum anderen kann der Rechtschreibprozess auf zwei unterschiedlichen Wegen der *Phonem-Graphem-Zuordnungen* erfolgen. Dies sind die komplexeren Prozesse. Auf dem ersten Weg der Phonem-Graphem-Zuordnung wird jedem Phonem entsprechend der Häufigkeit der wahrscheinlichen Zuordnung *direkt* ein Graphem zugeordnet. Dabei werden die gespeicherten Zuordnungen durch Wissen um Rechtschreibregeln ergänzt. Der zweite Weg der Phonem-Graphem-Zuordnung beruht auf einem *Erzeugungs- und Testprozess*

(generate-and-test process). Dabei werden mit einem *Rechtschreibgenerator* (spelling generator), der auf gespeicherten Phonem-Graphem-Korrespondenzen basiert, verschiedene Wortschreibungen erzeugt. Diese Probeschreibungen werden mit einem *Erkenner* (recognizer) auf gespeicherte visuelle Wortinformationen getestet und überprüft. Die Schreibungen werden als richtig bestätigt, wenn das Wortbild mit den gespeicherten Informationen des Erkenners übereinstimmt. Das Generator-Test-Prozessmodell ist in Abbildung 10 dargestellt.

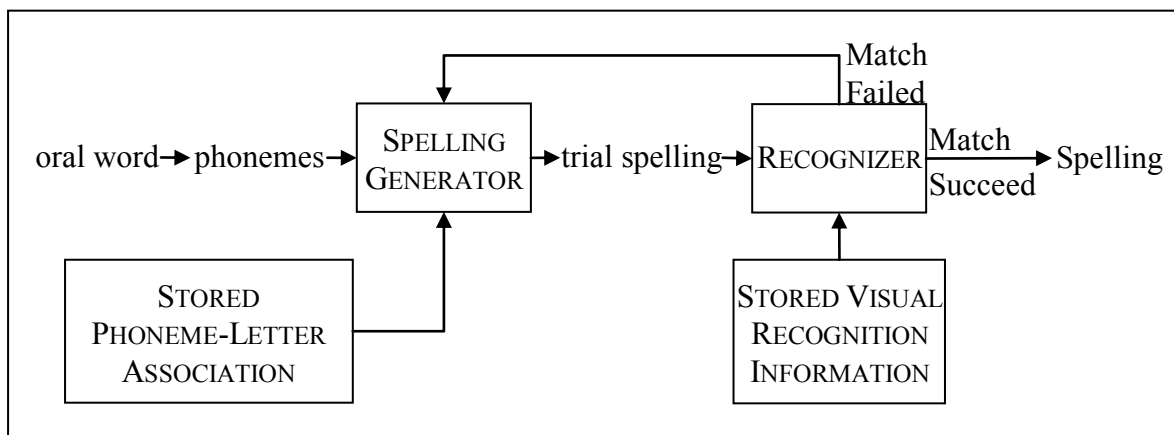


Abbildung 10: Spelling by means of a generator-test process (D. P. Simon & H. A. Simon, 1973, S. 118)

Der allgemeine Prozess des Rechtschreibens nach D. P. Simon und H. A. Simon (1973) kann so beschrieben werden, dass das zu schreibende Wort zunächst akustisch wahrgenommen wird. Ist dieses Wort vertraut und kann sich an die Schreibung erinnert werden, wird es aus dem Gedächtnis abgerufen und geschrieben. Kann sich nicht an die korrekte Wortschreibung erinnert werden, wird das Wort in Phoneme gegliedert. Im Rechtschreibgenerator werden dann den einzelnen Phonemen Grapheme zugeordnet. Anschließend wird das Wort geschrieben. Nun setzt der Erkenner ein und überprüft die Probeschreibung. Wird eine Übereinstimmung zwischen dem ganzen Wort und den gespeicherten visuellen Wortinformationen gefunden, ist der Rechtschreibprozess abgeschlossen. Findet sich keine Übereinstimmung, werden solange Wortschreibungen ausprobiert, bis es zu einer Übereinstimmung kommt.

In einer späteren Formulierung wird das Generator-Test-Prozessmodell durch zusätzliche Annahmen über die Ausnutzung von Gedächtniskapazitäten differenziert. Es wird angenommen, dass Wörter mit abweichender Phonem-Graphem-Korrespondenz markiert sind. Das bedeutet, dass bei Wortschreibungen, die einer besonderen Aufmerksamkeit

bedürfen, Gedächtnisstützen, Rechtschreibhilfen, Regeln sowie auditive, visuelle, motorische, semantische und morphologische Informationen genutzt werden. Des Weiteren wird angenommen, dass Wörter, die nicht aus dem Gedächtnis abgerufen werden können, nicht sofort in Phoneme zergliedert werden. Zunächst werden sie in größere sprachliche Einheiten wie Silben und Morpheme gegliedert. Erst wenn das Schreiben dieser sprachlichen Einheiten nicht möglich ist, wird auf Phonem-Graphem-Korrespondenzen zurückgegriffen, wobei bekannte Hilfen und Regeln verwendet werden. Die Schreibung wird anschließend lesend oder durch externe Gedächtnisstützen, wie z. B. ein Wörterbuch, auf Richtigkeit überprüft (D. P. Simon, 1976).

Das Generator-Test-Prozessmodell von D. P. Simon und H. A. Simon (1973) betont die Vielfalt unterschiedlicher Teilfertigkeiten, die für das Rechtschreiben notwendig sind. Nach dem Modell sind Kenntnisse über Phonem-Graphem-Korrespondenzen, die Fähigkeit, Wörter in sprachliche Einheiten gliedern zu können sowie Wissen um Rechtschreibregeln und verschiedene Wortschreibungen erforderlich (Klicpera et al., 2007). Die Annahme, Schreibungen werden mit visuellen Wortinformationen verglichen, ist heute allerdings überholt (Augst & Dehn, 2007). Hingegen wird bestätigt, dass beim Rechtschreiben zwei unterschiedliche Prozesse anzunehmen sind. Einerseits ist Rechtschreiben über das direkte Abrufen von Schreibungen aus einem inneren orthographischen Lexikon, andererseits durch das Konstruieren von Wörtern durch Phonem-Graphem-Korrespondenzen möglich (Klicpera & Gasteiger-Klicpera, 1993).

3.4.2 Zwei-Wege-Modelle des Rechtschreibens

D. P. Simon und H. A. Simon (1973) beschreiben in ihrem Modell den Rechtschreibprozess sehr ausführlich. Zwei-Wege-Modelle versuchen, nur die Generierung der Schreibung und nicht den Prüfprozess abzubilden. Sie sind daher keine unmittelbaren Nachfolgemodelle des Generator-Test-Prozessmodells (Mannhaupt, 2001).

3.4.2.1 Das Zwei-Wege-Modell von Barry (1994)

Das Zwei-Wege-Modell des Rechtschreibens von Barry (1994) beruht auf Ergebnissen

neuropsychologischer Studien mit dysgraphischen Erwachsenen, die durch neurologische Störungen, z. B. Schlaganfälle, Störungen der Schriftsprache erworben haben. Die Studien, oft Einzelfallbeschreibungen, verweisen auf eine Unterscheidung zwischen *Oberflächendysgraphie* und *phonologischer Dysgraphie*. Personen mit Oberflächendysgraphie können Pseudowörter und Wörter mit regelmäßiger, aber nicht mit unregelmäßiger Phonem-Graphem-Korrespondenz korrekt schreiben. Somit sind Störungen des Abrufs orthographischer Informationen zu vermuten. Personen mit phonologischer Dysgraphie können Pseudowörter nicht korrekt schreiben. Somit sind Störungen in der Zuordnung von Phonemen zu Graphemen zu vermuten. Diese Daten dienen der Bildung und Prüfung von Theorien über kognitive Verarbeitungsprozesse beim Rechtschreiben. Aufgrund der zwei in den Studien erfassten Störungen des Rechtschreibens, nimmt Barry folglich zwei voneinander unabhängige Prozesse an. Für das Rechtschreiben unterscheidet er zwischen einem *lexikalischen* (Lex) und einem *phonologischen* (Ass) Zugangsweg. Beim lexikalischen Zugangsweg werden bekannte Wortschreibungen aus dem *orthographischen Output-Lexikon* (orthographic output lexicon) abgerufen. Ein zu schreibendes Wort würde z. B. bei einem Diktat als bekannt registriert und die entsprechende Wortschreibung über die *Wortbedeutung* (L-sem), dessen *phonologische Repräsentation* (L-phon) oder *direkt* (L-dir?) aus dem Lexikon abgerufen werden. Beim phonologischen Zugangsweg werden unbekannte Wortschreibungen, z. B. Pseudowörter, aus dem *phonologischen Output-Lexikon* (phonological output lexicon) konstruiert. Ein zu schreibendes Wort würde z. B. bei einem Diktat als unbekannt registriert und über die Zuordnung von Phonemen und Graphemen konstruiert werden. Das Zwei-Wege-Modell des Rechtschreibens von Barry (1994) ist in Abbildung 11 dargestellt.

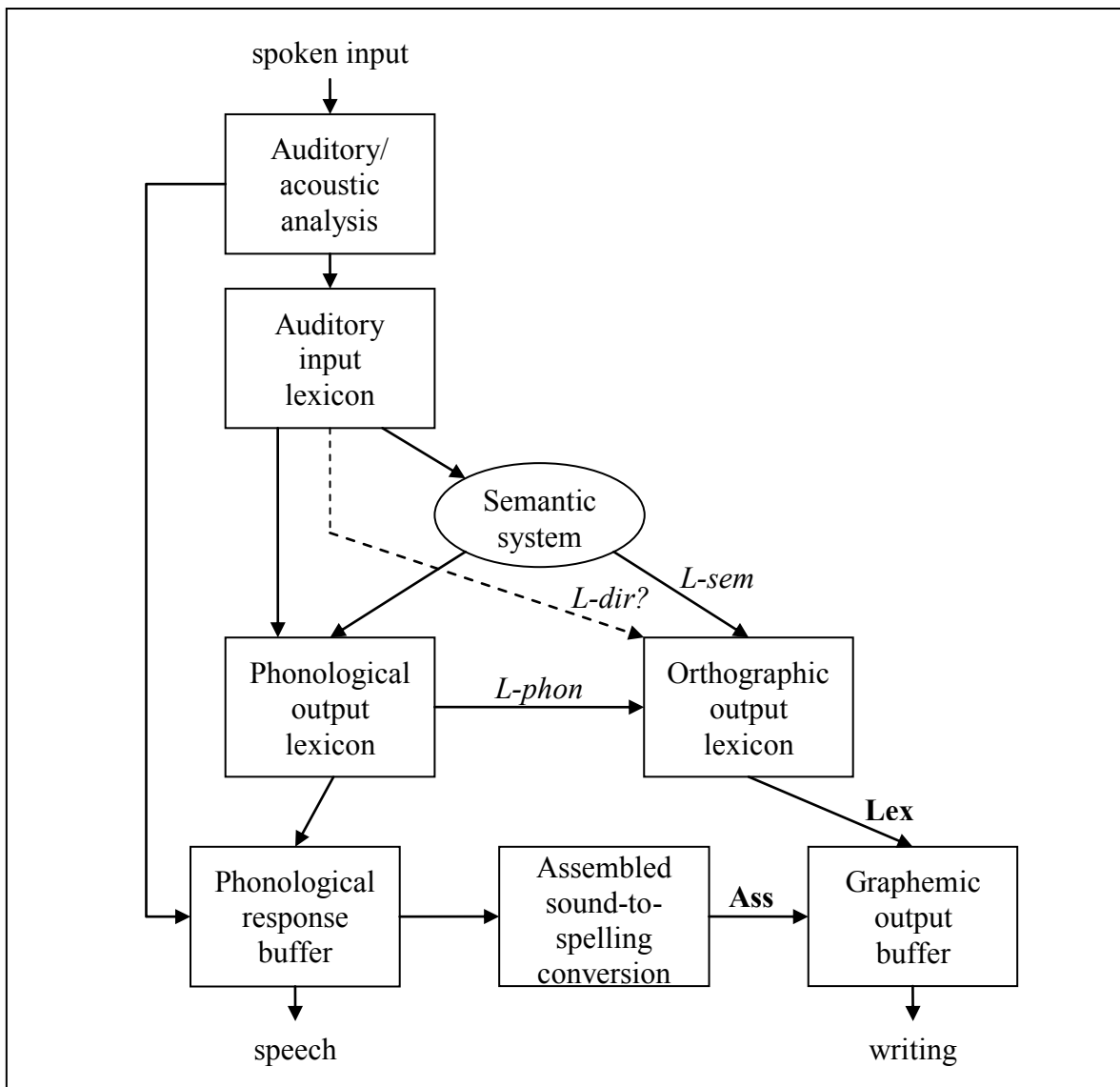


Abbildung 11: Routes to the orthographic output lexicon (Barry, 1994, S. 45)

Anmerkungen: Lex = lexical route, Ass = assembled route, L-sem = lexical semantic route, L-dir? = lexical direct route, L-phon = lexical phonological route

Nach Barry (1994) kann das Rechtschreiben über einen phonologischen und drei lexikalische Zugangswege erfolgen. Der allgemeine Rechtschreibprozess kann wie folgt beschrieben werden. Zunächst wird ein gesprochenes Wort durch die *auditiv-akustische Analyse* (auditory/acoustic analysis) perzeptiv erfasst, identifiziert und kategorisiert. Das akustische Signal wird wahrgenommen und in sprachliche und nichtsprachliche Reize unterschieden. Das als sprachlich wahrgenommene gesprochene Wort wird als eine Lautfolge identifiziert. Anschließend werden die identifizierten Phone einer Phonemkette zugeordnet. Dadurch können einzelne sprachliche Einheiten wie Wörter und Pausen im Sprechfluss voneinander unterschieden werden. Erst mit dem Übergang in das *auditive Input-Lexikon* (auditory input lexicon) werden die sprachlich identifizierten Einheiten aber als Wörter oder Pseudowörter verstanden, denn im auditiven Input-Lexikon werden die in

ihm gespeicherten phonologischen Repräsentationen mit den wahrgenommenen sprachlichen Reizen abgeglichen (Barry, 1994; Costard, 2007).

Ist die Wortschreibung der phonologischen Repräsentation bekannt, das heißt, ist ein Eintrag im orthographischen Output-Lexikon vorhanden, wird die Schreibung aus diesem abgerufen. Dieser Abruf erfolgt über drei unterschiedliche lexikalische Zugangswege. Die Routen greifen dabei auf verschiedene Wortinformationen aus Teilsystem zurück und aktivieren damit den Abruf aus dem orthographischen Output-Lexikon. Der lexikalisch-semanticale Weg (L-sem) greift auf Wortinformationen aus dem *semantic system* (semantic system) zurück, wodurch der phonologischen Repräsentation eine Wortbedeutung zugeordnet wird. Der Zugriff auf das semantische System aktiviert den Abruf aus dem orthographischen Output-Lexikon und führt zur graphematischen Repräsentation. Der lexikalisch-phonologische Weg (L-phon) greift direkt auf die phonologische Repräsentation im phonologischen Output-Lexikon zurück. Der Zugriff auf das phonologische Output-Lexikon aktiviert den Abruf aus dem orthographischen Output-Lexikon und führt zur graphematischen Repräsentation. Der direkte lexikalische Weg (L-dir?) greift unmittelbar auf das orthographische Output-Lexikon zurück, aktiviert den Abruf und führt zur graphematischen Repräsentation. Gibt es Wörter mit gleicher phonologischer Repräsentation, sogenannte Homophone, wird die Schreibung über den lexikalisch-semanticalen (L-sem) und lexikalisch-phonologischen (L-phon) Weg abgerufen. Der lexikalisch-semanticale Weg (L-sem) greift zunächst über das semantische System auf Kontextinformationen und Wortbedeutungen der Homophone zurück. Das semantische System ordnet diese Wortinformationen der phonologischen Repräsentation im phonologischen Output-Lexikon zu. Der Zugriff auf das phonologische Output-Lexikon aktiviert den Abruf aus dem orthographischen Output-Lexikon und führt zu eindeutigen graphematischen Repräsentationen (Barry, 1994; Cholewa, 2004).

Ist die Wortschreibung der phonologischen Repräsentation unbekannt, das heißt, ist kein Eintrag im orthographischen Output-Lexikon vorhanden, wird die Schreibung über den phonologischen Zugangsweg konstruiert. Dabei wird die phonologische Repräsentation zunächst in einem *phonologischen Antwort-Buffer* (phonological response buffer) kurzfristig gespeichert, um in einem nächsten Schritt die phonologische Repräsentation in Phoneme segmentieren und ihnen Grapheme zuordnen zu können (assembled sound-to-spelling conversion). Das Ergebnis ist eine graphematische Repräsentation. Nach dem

lexikalischen oder phonologischen Verarbeitungsprozess wird die graphematische Repräsentation in einem *graphematischen Antwort-Buffer* (graphematic response buffer) kurzfristig aufrechterhalten, um diese anschließend schreiben zu können (Barry, 1994; Cholewa, 2004; Costard, 2007).

Barry (1994.) resümiert, dass „... the dual-route model of spelling has been supported most persuasively by the erroneous spelling performance of patients with disorders of spelling competence, and in particular by the dissociations observed within patients with phonological and surface dysgraphia“ (S. 36). Demzufolge können nur solche Verarbeitungsprozesse beim Rechtschreiben angenommen werden, die sich durch Störungen der Schriftsprache bei Personen mit erworbenen Hirnläsionen zeigen. Daher verwundert es nicht, dass bei neu zu beobachtenden Störungen der Schriftsprache, weitere neue Prozesse angenommen werden müssen (Mannhaupt, 2001). Barry (1994) selbst hinterfragt sein Zwei-Wege-Modell kritisch, indem er anführt, dass die phonologische und lexikalische Route miteinander interagieren. Dies geht aus seinem Modell jedoch nicht hervor. „Simply stated ... the point of this integration is to allow sublexical information to strengthen targeted lexical representations. Thus, models of spelling must have a mechanism by which sublexical activation can interact with and strengthen correct lexical responses ...“ (Folk, Rapp & Goldrick, 2002, S. 665). Rapp, Epstein und Tainturier (2002) z. B. entwickelten ein entsprechendes Modell. Sie nehmen an, dass die aus der phonologischen und lexikalischen Route gewonnenen Wortinformationen in einer sogenannten *Graphemebene* (graphem layer) zusammengeführt werden. Wird ein zu schreibendes Wort akustisch wahrgenommen, werden beide Zugangswege aktiviert, um aus der Graphemebene graphematische Informationen abzurufen. Rapp et al. nehmen des Weiteren an, dass es einen *Rücklauf* (feedback) von aktivierten graphematischen Informationen der Graphemebene zu gespeicherten orthographischen Wortformen im orthographischen Output-Lexikon gibt. Die durch die beiden Zugangswege aktivierten Wortinformationen eines bestimmten Wortes in der Graphemebene werden zusammengefasst. Sie bilden eine Summe. Die summierten Wortinformationen werden durch den Rücklauf dem orthographischen Output-Lexikon gegenübergestellt. Durch die summierten und aktivierten phonologischen und lexikalischen Wortinformationen kann die graphematische Wortform ausgewählt werden.

3.4.2.2 Das Zwei-Komponenten-Modell von Scheerer-Neumann (2003)

Scheerer-Neumann (2003) entwirft ein Prozessmodell des Rechtschreibens, das von zwei voneinander getrennten Gedächtnisspeichern ausgeht. Es wird angenommen, dass Rechtschreiben auf der Grundlage von *Regelwissen* und einem *inneren orthographischen Lexikon* erfolgt (Scheerer-Neumann, 1987). Das Regelwissen umfasst gespeicherte Informationen zu schriftsprachlichen Regelmäßigkeiten. Dazu zählen Kenntnisse über Phonem-Graphem-Korrespondenzen und über ihre Zuordnungshäufigkeiten. Das Regelwissen umfasst zudem Wissen über orthographische Regelmäßigkeiten. Dazu zählen Kenntnisse über legale und illegale Graphemfolgen sowie Wortbildungsregeln. Das innere orthographische Lexikon speichert unterschiedlich miteinander verknüpfte wortspezifische Informationen. Das innere orthographische Lexikon wird in ein Zentrum und in eine Peripherie unterteilt. Im Zentrum werden die Grapheme eines Wortes und ihre dazugehörigen Graphem-Namen, Graphem-Lautwerte, ihre visuellen und schreibmotorischen Merkmale gespeichert. Sie werden dort auch als abstrakte Repräsentationen gespeichert. In der Peripherie werden weitere integrale Worteigenschaften wie die Bedeutung und der emotionale Wert, die phonologische Repräsentation, orthographische Regelmäßigkeiten, der morphologische Aufbau sowie visuelle und schreibmotorische Merkmale des ganzen Wortes gespeichert (Scheerer-Neumann, 1986, 1987). Abbildung 12 gibt die zwei Gedächtnisspeicher wieder.

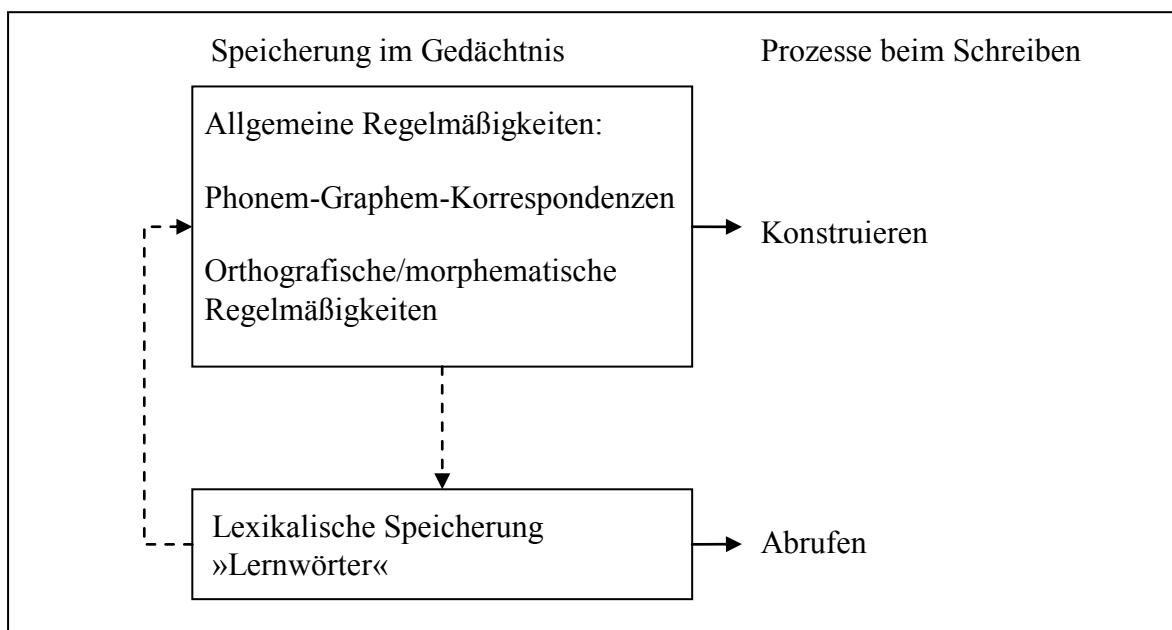


Abbildung 12: Zwei-Komponenten-Modell der Rechtschreibung (Scheerer-Neumann, 2003, S. 51)

Der allgemeine Prozess des Rechtschreibens nach Scheerer-Neumann (2003) kann so beschrieben werden, dass beim Schreiben eines unbekannten Wortes die Schreibung aus den gespeicherten schriftsprachlichen Regelmäßigkeiten *konstruiert* wird. Dabei werden bereits gespeicherte Phonem-Graphem-Korrespondenz, orthographische Regelmäßigkeiten und Wortbildungsregeln berücksichtigt. Beim Schreiben eines bekannten Wortes werden gespeicherte wortspezifische Informationen, so genannte *Lernwörter*, aus dem inneren orthographischen Lexikon *abgerufen*. Dabei kann auf alle Lernwörter zurückgegriffen werden, die bisher wortspezifisch geübt und gespeichert wurden. Das bedeutet aber nicht, dass alle Lernwörter schon orthographisch korrekt gespeichert sind (Scheerer-Neumann, 2003, 2006). Dass beide Speichersysteme nicht unabhängig voneinander existieren, zeigt das Ineinandergreifen beider Systeme während der Entwicklung des Rechtschreibens. Zu Beginn der Rechtschreibentwicklung sind keine Einsichten in die Regelmäßigkeiten der Schriftsprache vorhanden. Erste Wörter können daher nur über deren phonologische Repräsentation und Bedeutung mit einer Graphemfolge assoziiert und im inneren orthographischen Lexikon gespeichert werden. Die Grapheme werden beim beginnenden Rechtschreiben somit abgerufen und nicht konstruiert. Erst mit beginnenden Einsichten in die Phonem-Graphem-Korrespondenzen können Wörter konstruiert werden. Erste wortspezifische Speicherungen im inneren orthographischen Lexikon sind anzunehmen, wenn Kinder Wortschreibungen kennen, die von einfachen Phonem-Graphem-Korrespondenzen abweichen. Zu diesem Zeitpunkt werden alphabetische Konstruktionen mit wortspezifischen Abrufungen kombiniert. Mit wachsendem Wissen um orthographische Regelmäßigkeiten werden wortspezifische Schwierigkeiten für Kinder verständlich und damit leichter einzuprägen. Gleichzeitig kann durch die Speicherung vieler Lernwörter das Verständnis für Regelmäßigkeiten erleichtert werden (Scheerer-Neumann, 1987, 2003). „Im Laufe des Erwerbs der Rechtschreibung verändert sich beim Schreiben das Verhältnis zwischen eigenaktiven Konstruktionen und dem Abrufen von Lernwörtern bzw. der sie konstituierenden Morpheme zugunsten des Abrufprozesses. Allerdings bleiben phonographische Prozesse zur Handlungssteuerung auch beim Erwachsenen erhalten . . .“ (Scheerer-Neumann, 2008, S. 167).

Der Verdienst des Zwei-Komponenten-Modells der Rechtschreibung von Scheerer-Neumann (2003) ist, dass im Rahmen des Prozessmodells die Rechtschreibentwicklung abgebildet werden kann. Beide Gedächtnisspeicher verarbeiten Schriftsprache entsprechend der jeweiligen Stufe qualitativ unterschiedlich. Durch das Zwei-Wege-

Modell von Scheerer-Neumann (2003) kann hinter jeder Entwicklungsstufe des Rechtschreiberwerbs ein anders ablaufender Teilprozess des Rechtschreibens angenommen werden. Gleichzeitig wird dadurch das Regellernen und wortspezifische Lernen hervorgehoben. Sowohl Regelwissen als auch wortspezifische Informationen sollen im Unterricht vermittelt werden. Dabei ist die individuelle Rechtschreibentwicklung zu beachten. Für das zweite Schuljahr bedeutet das, dass es keinen Sinn hat, eine sehr große Anzahl von Lernwörtern einzuprägen, wenn beispielsweise kritische Phonem-Graphem-Korrespondenzen oder morphematische Zusammenhänge noch nicht erkannt werden, denn durch das noch unvollständige Wissen um orthographische Regelmäßigkeiten würde das Einprägen vieler Lernwörter zu diesem Zeitpunkt einen enormen Lernaufwand bedeuten (Scheerer-Neumann, 2006). Ein weiterer, bislang unberücksichtigter Verdienst des Zwei-Komponenten-Modells von Scheerer-Neumann (2003) liegt in der Annahme des Speichers für vielfältiges schriftsprachliches Regelwissen.

„Diese erweiterte Sicht des ‚Regelspeichers‘ macht es auch möglich zu erklären, warum kompetente Rechtschreiber Kunstwörter (die sie also noch nie gesehen haben) zumindest dann den orthographischen Regeln entsprechend schreiben können, wenn sie in einen Text eingebettet sind, wie im folgenden Beispiel (Menzel 1985, 26): ‚Novemberwetter Die /fértenen/ Wolken /tü:len/ über das Land.Es [sic] /bläkst/ der Sturm im Gebälk.‘ Selbst wenn die SchreiberInnen nicht angeben können, warum sie ein Kunstwort auf eine bestimmte Weise (richtig) geschrieben haben, kann angenommen werden, dass hier in einem Hypothesenbildungsprozess Regelwissen und Analogien zu gespeicherten Wörtern benutzt wurden, um die ‚korrekte‘ Schreibung zu finden: z.B. dass ‚tü:len‘ ein Verb sein muss und deshalb nicht groß geschrieben werden kann usw.“ (Richter, 2007, S. 157).

3.4.3 Netzwerkmodelle

Netzwerkmodelle nehmen nicht wie Zwei-Wege-Modelle Modelle ein inneres orthographisches Lexikon und zwei unterschiedliche Verarbeitungsprozesse beim Rechtschreiben an, sondern in diesen Modellen wird die Annahme vertreten, dass die für das Rechtschreiben notwendigen Informationen in Netzwerken gespeichert und durch vieles Üben und Wiederholen über Zuordnungen miteinander verbunden sind. Netzwerkmodelle sind Computersimulationsmodelle (Costard, 2007; Klicpera et al., 2007).

3.4.3.1 Das Netzwerkmodell des Rechtschreibens von Brown und Loosemore (1994)

Das Netzwerkmodell des Rechtschreibens von Brown und Loosemore (1994) beruht auf Erkenntnissen über die menschliche Gehirnstruktur. Kurz, das Gehirn besteht aus Milliarden von Nervenzellen, die über Nervenbahnen miteinander verbunden sind. Netzwerkmodelle versuchen, dieses Nervennetz zu simulieren, um kognitive Prozesse wie z. B. das Rechtschreiben aufzuklären. Netzwerkmodelle bestehen aus künstlichen Neuronen, die Einheiten darstellen und mit unterschiedlicher Intensität miteinander verbunden sind. Diese Neuronen besitzen einen *Aktivierungsgrad* (activation value), der angibt, ob sich die Neuronen gegenseitig hemmen oder erregen. Die Stärke des Aktivierungsgrades ist abhängig von der Intensität der neuronalen Verbindungen. Das bedeutet, je stärker die neuronalen Verbindungen zwischen den Einheiten durch die jeweilige Verwendungsfrequenz sind, desto besser führt ein bestimmtes Aktivierungsmuster zur Aktivierung aller mit einer Einheit verbundenen Neuronen (Brown & Loosemore, 1994).

Das Rechtschreibprozessmodell von Brown und Loosemore (1994) ist entsprechend der allgemeinen Struktur von Netzwerkmodellen aufgebaut. Es besteht aus drei Einheiten, den *Output-Einheiten* (output units), den *verborgenen Einheiten* (hidden units) und den *Input-Einheiten* (input units). Die Output-Einheiten umfassen die graphematischen und die Input-Einheiten die phonologischen Wortformen. Die verborgenen Einheiten verbinden die Output- und Input-Einheiten so miteinander, dass keine direkte Beziehung zwischen ihnen besteht. Je mehr verborgene Einheiten vorhanden sind, desto größer sind die Anzahl der Verbindungen und die Fähigkeit, neue Assoziationen zu lernen. Das Netzwerkmodell des Rechtschreibens von Brown und Loosemore ist in Abbildung 13 veranschaulicht.

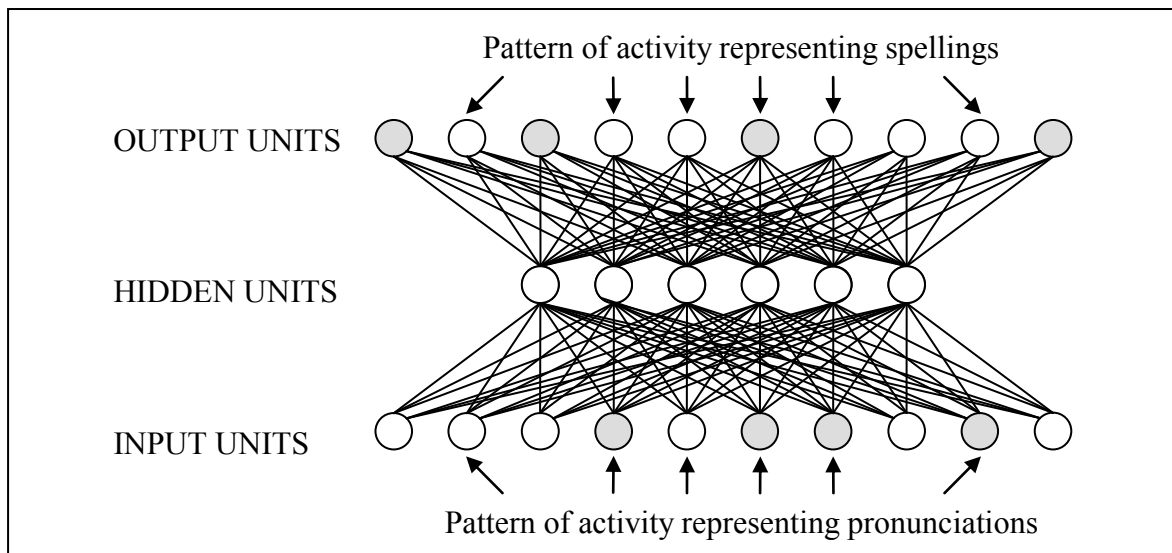


Abbildung 13: The architecture of the connectionist model. Only some units and connections are illustrated. (Brown & Loosemore, 1994, S. 323)

Der allgemeine Prozess des Rechtschreibens basiert nach Brown und Loosemore (1994) auf einem Lernalgorithmus. Bei Schreibanfängern sind die Aktivierungsmuster und Verbindungen zwischen den phonologischen und graphematischen Wortformen noch nicht gefestigt und erfolgen daher zufällig. Durch wiederholende Lernprozesse nehmen der Aktivierungsgrad und die Verbindungsstärke zu. Infolgedessen kommt es zu einer schrittweisen Annäherung an die Verbindungen, die die korrekte Schreibweise repräsentieren. Fehlschreibungen treten seltener auf.

Brown und Loosemore (1994) beleuchten den Rechtschreibprozess mit ihrem Modell auf eine neue Art, „as a task of mastering the statistical associations between a set of patterns representing the phonological forms of words and a set of patterns representing the orthographic forms“ (S. 333). Für die empirische Überprüfung des Rechtschreibprozessmodells haben Brown und Loosemore (1994) Schreibungen von Kindern unterschiedlicher Rechtschreibfertigkeiten mit denen eines auf diesem Modell basierenden Computerprogramms verglichen. Die Ergebnisse zeigen, dass die Fehleranzahl sowohl bei den Kindern als auch beim Computerprogramm mit steigender Anzahl von Lernzyklen sinkt. Zudem zeigen sich zwischen den Kindern und dem Computerprogramm Übereinstimmungen in der Schwierigkeit von Wortschreibungen. Auch bei einem Vergleich zwischen Kindern mit Rechtschreibschwierigkeiten und einer veränderten Version des Computerprogramms, bei dem die Anzahl der verborgenen Einheiten reduziert wurde, zeigen sich Übereinstimmungen in der Fehleranzahl (Brown & Loosemore, 1994). „Sie [Brown und Loosemore] kommen deshalb zu dem Schluß [sic], daß [sic] ihr

konnektionistisches Modell sowohl den normalen als auch den Rechtschreibprozeß [sic] bei Lese-Rechtschreibschwäche simulieren und damit beschreiben kann“ (Mannhaupt, 2001, S. 58).

3.4.3.2 Das Rechtschreibprozessmodell NETspell von Olson und Caramazza (1994)

Das Rechtschreibprozessmodell *NETspell* von Olson und Caramazza (1994) ist ein Netzwerkmodell, das im Vergleich zu anderen Netzwerkmodellen *lexikalische* und *suprasegmentale Informationen* mit berücksichtigt. Es beruht auf Erkenntnissen über Auftretenswahrscheinlichkeiten von Phonem-Graphem-Korrespondenzen in unterschiedlichen Wortpositionen und Studien mit gehirnerkrankten Personen. Zum Ersten, nicht jedes Graphem kommt in allen Wortpositionen vor. Netzwerke müssen diese Informationen nicht auswendig lernen, sondern sind in der Lage, sie zu verallgemeinern. Zum Zweiten, neuropsychologische Studien belegen, dass gehirnerkrankte Personen beim Schreiben typische Fehler zeigen, die auf Ausfälle lexikalischer bzw. nichtlexikalischer (suprasegmentale) Prozesse zurückzuführen sind. Somit sind lexikalische und suprasegmentale Prozesse beim Rechtschreibprozess beteiligt.

Das Netzwerkmodell von Olson und Caramazza (1994) versucht, den Rechtschreibprozess zu simulieren. Es besteht aus drei Ebenen, einer *Input-Ebene* (input layer), einer *Output-Ebene* (output layer) und einer *vermittelnden Ebene* (intermediate layer) mit *verborgenen Einheiten* (hidden units). Die Input-Ebene führt die Eingabe, die Phoneme eines zu schreibenden Wortes, über sieben *Phonemfenster* zum Netzwerk. Ein Phonemfenster speichert ein zentrales Phonem Z. Die anderen Phonemfenster speichern jeweils drei Phoneme auf der linken (3 2 1) und rechten (1 2 3) Seite. Sie stützen als zusätzliche Kontextinformationen die Verarbeitung. Die Output-Ebene kodiert die Ausgabe, die Grapheme des zu schreibenden Wortes. Die verborgenen Einheiten unterstützen das Überführen vom Input zum Output (Mannhaupt, 2001; Olson & Caramazza, 1994). Das Netzwerkmodell ist in Abbildung 14 in Anlehnung an die theoretischen Ausführungen von Olson und Caramazza (1994) sowie an eine Darstellung von Klicpera et al. (2007) dargestellt.

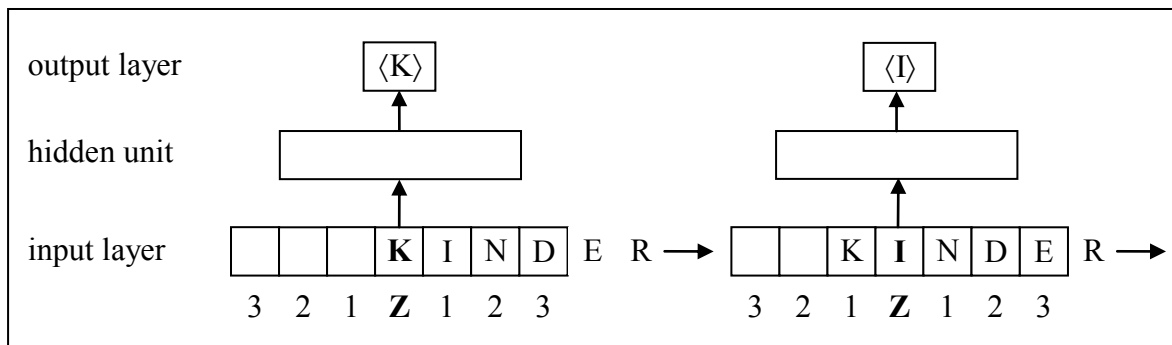


Abbildung 14: Rechtschreibprozessmodell NETspell (in Anlehnung an Klicpera et al., 2007; Olson & Caramazza, 1994)

Der allgemeine Rechtschreibprozess kann nach dem Modell von Olson und Caramazza (1994) wie folgt beschrieben werden. Zunächst werden das erste Phonem eines zu schreibenden Wortes in das zentrale Z und die folgenden Phoneme in die rechten (1 2 3) Phonemfenster der Input-Ebene eingeordnet. Dann werden die Phoneme von rechts nach links bewegt. Somit nimmt jedes Phonem einmal die zentrale Position ein, wodurch jedem Phonem über die mittlere Ebene in der Output-Ebene ein Graphem zugeordnet wird. Die Zuordnung von Phonemen zu Graphemen erfolgt über die Verbindungen zwischen der Input-Ebene und den verborgenen Einheiten der vermittelnden Ebene. Die vermittelnde Ebene ist auch mit der Output-Ebene verbunden. Die Verbindungsstärke zwischen den einzelnen Ebenen bestimmt den Aktivierungsgrad, das heißt, wie stark eine Einheit durch die vorhergehende aktiviert wird. Ist ein bestimmter Aktivierungsgrad erreicht, werden die Einheiten der Output-Ebene eingeschaltet. Das bedeutet, dass die aktivierte, eingeschaltete Output-Ebene für ein Phonem die jeweilige Schreibweise vermutet und dem System präsentiert. Ist die Vermutung falsch, wird die Verbindungsstärke für das falsch zugeordnete Graphem so geändert, dass das Netzwerk in einem nächsten Durchgang bei gleichem Input-Output-Paar die richtige Antwort hervorbringt. Dieser Vorgang wird immer wiederholt und führt damit zu einer höheren Verbindungsstärke und zu genaueren Antworten (Klicpera et al., 2007; Olson & Caramazza, 1994).

In einem Simulationstraining des Netzwerkmodells berichten Olson und Caramazza (1994), dass NETspell in 60 Durchgängen 1000 bzw. 1628 Wörter schreiben lernt. Auf der Ebene der Phonem-Graphem-Korrespondenzen erreicht das Modell 94 % bzw. 96 % Richtigschreibungen. Auf der Wortebene erreicht das Modell 70 % bzw. 83 % Richtigschreibungen. Fehlschreibungen basierten auf Regularisierungsfehlern, das heißt Wörter mit unregelmäßiger Phonem-Graphem-Korrespondenz wurden regelmäßig

geschrieben, und auf Nullreaktionen, bei denen keine Zuordnung von Phonemen und Graphemen stattfand. Für die weitere Überprüfung haben Olson und Caramazza (1994) dem Modell Läsionen hinzugefügt, um Fehlermuster des Programms mit denen gehirnkranker Personen zu vergleichen. Dabei zeigte sich, dass das Programm ähnliche Fehler wie gehirnkranke Personen produziert. Folglich kann es einen Großteil von Schreibungen simulieren.

Bei genauerer Betrachtung werden Probleme des Netzwerkmodells von Olson und Caramazza (1994) deutlich. Das Modell ist z. B. nicht in der Lage, zu 100 % alle Wörter richtig zu schreiben. Dies gelingt auch bei Veränderungen des Modells, z. B. Erhöhungen der Durchläufe und der Anzahl verborgener Einheiten, nicht. Olson und Caramazza (1994) können die Fehlerursache nicht benennen. Sie vermuten aber, dass die Berücksichtigung zusätzlicher, orthographischer Informationen zu einer erhöhten Anzahl von Richtigschreibungen führen würde. Olson und Caramazza (1994) bewerten ihr Netzwerkmodell ähnlich wie Brown und Loosemore (1994) nicht als ausgereifte Theorie, sondern als Auftakt zu einer umfassenden Forschungstätigkeit.

3.4.4 Zusammenfassung und Schlussfolgerungen

Der Überblick über die hier dargestellten Modelle des Rechtschreibprozesses verdeutlicht, dass ausgehend vom Generator-Test-Prozessmodell (D. P. Simon & H. A. Simon, 1973) im Wesentlichen zwei Arten von Rechtschreibprozessmodellen entwickelt wurden: Zwei-Wege-Modelle und Netzwerkmodelle. Zwei-Wege-Modelle gehen davon aus, dass Rechtschreiben über zwei miteinander agierende Wege erfolgt. Beim ersten Weg, dem direkten Zugriff, wird das Wort aus dem mentalen Lexikon abgerufen. Beim zweiten Weg, dem indirekten Zugriff, wird das Wort über Phonem-Graphem-Korrespondenzen und dem Wissen um orthographische Regelmäßigkeiten konstruiert. Netzwerkmodelle hingegen verzichten auf die Annahme zweier Verarbeitungsprozesse beim Rechtschreiben. Sie gehen davon aus, dass der Rechtschreibprozess auf Zuordnungswahrscheinlichkeiten basiert (Costard, 2007; Klicpera et al., 2007).

Alle hier dargestellten Prozessmodelle beschreiben die kognitiven Prozesse, die bei einem *geübten Rechtschreiber* von einer phonetischen Repräsentation zu einer orthographisch

korrekten Repräsentation führen. Allein die Zwei-Wege-Modelle lassen Rechtschreibwege und -prozesse erkennen, auf die *Rechtschreibanfänger* zugreifen können. Wie schreiben nun also Rechtschreibanfänger? Zuerst nimmt das Kind ein akustisches Signal wahr und identifiziert es als ein gesprochenes Wort, z. B. [li:t]. Anschließend versucht das Kind, die Laute schrittweise herauszuhören und zu verschriften. Dabei wird in einem ersten Schritt der Anfangslaut analysiert, meist durch leises, überdeutliches Sprechen vom ganzen Wort abgegrenzt; [l]. Danach muss das Kind überlegen, welches Graphem zu dem Phonem passt, das schreibmotorische Programm abrufen und das Graphem schreiben; ⟨L⟩. Während dieser Teilschritte wird der Rest des Wortes im Kurzzeitgedächtnis für die nächste Lautanalyse solange präsent gehalten bis der Rechtschreibvorgang für das Graphem abgeschlossen ist. Im zweiten Schritt versucht das Kind, den zweiten Laut herauszuhören und zu verschriften. Dabei, und bei jedem weiteren Laut, wiederholt sich dieser Prozess: „Anlautanalyse, Graphemzuordnung, Verschriftlichung des Graphems, Vergegenwärtigen – Nachhalten des Geschriebenen“ (Mannhaupt, 2010, S. 188). Wird diese Prozessabfolge vom Kind korrekt durchgeführt, kann das Kind das Wort ⟨LIT⟩ lautorientiert richtig schreiben. Sollte das Wort orthographisch richtig geschrieben werden, müsste die Prozessabfolge um orthographisches Regelwissen ergänzt werden, sodass das Kind das Wort *Lied* orthographisch richtig schreibt. Der abschließende Schritt im Rechtschreibprozess ist die Kontrolle des Geschriebenen. Sie bildet sich aber erst in der 2. Klasse heraus, weil Kinder erst dann so sicher lesen können und einen entsprechenden Rechtschreibwortschatz haben, um falsch geschriebene Wörter auch als falsch zu identifizieren (Mannhaupt, 2010).

Zusammenfassend liegt der Nutzen von Rechtschreibprozessmodellen in der Feststellung, welche Fertigkeiten für das Rechtschreiben notwendig sind. In ihrer Gesamtheit betonen alle hier dargestellten Modelle vor allem folgende Komponenten beim Rechtschreiben:

1. die phonologische Bewusstheit als Fähigkeit, Aufmerksamkeit auf formale Aspekte der Sprache zu lenken,
2. die Fähigkeit, eine Graphemfolge während des Schreibens im Arbeitsgedächtnis möglichst lange zu behalten,
3. wortspezifische Kenntnisse für den lexikalischen Abruf von Wörtern,
4. Wissen um Ableitungsregeln und Rechtschreibregeln sowie
5. semantische Informationen, um die richtige Schreibweise z. B. für Homophone zu erschließen (Klicpera et al., 2007).

Aus den Ausführungen zu den Prozessmodellen ergibt sich für die Entwicklung der FE-RS 2 Folgendes: Für das dieser Arbeit zugrunde liegende Modell des Rechtschreibprozesses soll erneut die im Abschnitt 3.2.2 erläuterte Unterscheidung von Graphematik und Orthographie (Neef, 2005b) Berücksichtigung finden. Dieser Unterscheidung entspricht vor dem Hintergrund des Stufenmodells von Scheerer-Neumann (2008) auch ihr Zwei-Komponenten-Modell (2003), weshalb es hier als Grundlage genommen wird. Sowohl Richter (2007) als auch Jansen (2008) verweisen auf das pädagogisch-diagnostische Potenzial von Zwei-Wege-Modellen. „Bei der Erfassung von Rechtschreibleistungen bieten die Modelle zumindest potenziell die Möglichkeit, aus ihnen Variablen abzuleiten, die für die anwendungsbezogene Arbeit von hohem Informationswert sind“ (Jansen, 2008, S. 183). Vor allem die Frage nach dem Weg zur Schreibung, ob Regeln angewendet werden oder Schreibungen auf direktem Abruf basieren, scheint dabei wichtig zu sein, denn auf Kinder, die über viele Zwischenschritte Schreibungen konstruieren, muss anders eingegangen werden als auf Kinder, die ohne Zwischenschritte Schreibungen aus dem Gedächtnis abrufen (Jansen, 2008). Für die Entwicklung eines diagnostischen Verfahrens bedeutet das, dass „dann als mögliche Prüfitems solche Wörter in Frage [kommen], die das Kind noch nicht kennt (seltene Wörter) oder nicht kennen kann (Pseudowörter). Denn man würde erwarten, dass bei einer sich generalisierenden Regelanwendung auch Pseudowörter im Sinne der ‚Regel‘ korrekt geschrieben würden“ (Jansen, 2008, S. 183f.). Jansens (2008) Vorschlag soll an einem Beispiel überprüft werden: Für das Pseudowort [ra:l] gibt es drei mögliche graphematische Schreibungen: ⟨ral⟩, ⟨raal⟩, ⟨rahl⟩. Die Schreibung ⟨ral⟩ deutet allein auf eine Beherrschung der Phonem-Graphem-Korrespondenz hin. Die Schreibungen ⟨raal⟩ und ⟨rahl⟩ deuten auf die Anwendung des Dehnungsschreibens hin. Ob diese aber durch Anwendung einer orthographischen Regel oder durch Analogiebildung zu einem im orthographischen Lexikon gespeicherten Wort zustande gekommen ist, kann nicht gesagt werden. Ein Rückschluss, ob die Regel des Dehnungsschreibens beherrscht wird, ist folglich nicht möglich. Es sollte aber ein Anliegen der zukünftigen Forschung sein, theoriebasierte Instrumente zu entwickeln, die Lehrkräften Einblicke in die Rechtschreibprozesse der Kinder ermöglichen. Rechtschreibdiagnostik und -förderung könnten damit fundierter gestaltet werden. Da dieser Forschungsschritt jedoch noch aussteht, soll bei dem zu entwickelnden Verfahren nicht der Anspruch erhoben werden, Rechtschreibprozesse zu erfassen.

Da angenommen wird, dass sich der Rechtschreiberwerb auch über die Vermittlung im Unterricht ableiten lässt, werden im folgenden Abschnitt die Anforderungen und Inhalte des Lernbereichs Rechtschreiben in den Bildungsstandards im Fach Deutsch für den Primarbereich (KMK, 2005), in dem Rahmenplan Grundschule Deutsch von Brandenburg, Berlin, Bremen und Mecklenburg-Vorpommern (MBS et al., 2004) und in den zugelassenen Sprachbüchern aus dem Schulbuchkatalog 2009/2010 für das Unterrichtsfach Deutsch der Jahrgangsstufe 2 in Grundschulen in Mecklenburg-Vorpommern (MBWK, 2009a) beschrieben.

3.5 Anforderungen und Inhalte des Lernbereichs Rechtschreiben in den Bildungsstandards der KMK (2005), dem Rahmenplan von Mecklenburg-Vorpommern (2009a) und den Sprachbüchern

Bildungsstandards und Rahmenpläne sind verbindliche Vorgaben. Sie legen fest, welche Kompetenzen Schüler am Ende einer bestimmten Klassenstufe in den Kernaufgaben der Unterrichtsfächer erworben haben sollten. Um die Rechtschreibleistungen von Kindern mit einem diagnostischen Verfahren adäquat erfassen zu können, müssen die Anforderungen und Inhalte des Lernbereichs Rechtschreiben in den Bildungsstandards, dem Rahmenplan sowie den gängigen Schulbüchern berücksichtigt werden, denn es können nur die Rechtschreibleistungen erfasst und vorausgesetzt werden, welche auch gelehrt wurden.

3.5.1 Rechtschreiben in den Bildungsstandards der KMK (2005)

Die Bildungsstandards im Fach Deutsch für den Primarbereich (KMK, 2005) sind deutschlandweit für die vierte Klassenstufe beschlossen. Sie greifen allgemeine Bildungsziele auf und legen fest, welche fachbezogenen Kompetenzen am Ende der 4. Klasse in den Kernbereichen des Unterrichtsfaches Deutsch erwartet werden können. Gleichzeitig betonen sie die schulische und bildungspolitische Verantwortung für das Erreichen der Kompetenzen (Klieme et al., 2007).

Die Bildungsstandards Deutsch (KMK, 2005) unterscheiden vier Kompetenzbereiche voneinander. Den Kompetenzbereichen werden fachliche Schwerpunkte zugeordnet und

diesen wiederum die Standards. Rechtschreiben ist ein fachlicher Schwerpunkt des Kompetenzbereichs Schreiben. Zur orthographischen Kompetenz am Ende der 4. Klassenstufe heißt es:

„Die Kinder verfügen über grundlegende Rechtschreibstrategien. Sie können lautensprechend verschriften und berücksichtigen orthographische und morphematische Regelungen und grammatisches Wissen. Sie haben erste Einsichten in die Prinzipien der Rechtschreibung gewonnen. Sie erproben und vergleichen Schreibweisen und denken über sie nach. Sie gelangen durch Vergleichen, Nachschlagen im Wörterbuch und Anwenden von Regeln zur richtigen Schreibweise. Sie entwickeln Rechtschreibgespür und Selbstverantwortung ihren Texten gegenüber“ (KMK, 2005, S. 8).

Dem Schwerpunkt „richtig schreiben“ werden folgende Standards zugeordnet:

- „geübte, rechtschreibwichtige Wörter normgerecht schreiben,
- Rechtschreibstrategien verwenden: Mitsprechen, Ableiten, Einprägen,
- Zeichensetzung beachten: Punkt, Fragezeichen, Ausrufezeichen, Zeichen bei wörtlicher Rede,
- über Fehlersensibilität und Rechtschreibgespür verfügen
- Rechtschreibhilfen verwenden
 - Wörterbuch nutzen,
 - Rechtschreibhilfen des Computers kritisch nutzen
- Arbeitstechniken nutzen
 - methodisch sinnvoll abschreiben
 - Übungsformen selbstständig nutzen,
 - Texte auf orthographische Richtigkeit überprüfen und korrigieren“ (KMK, 2005, S. 10f.).

3.5.2 Rechtschreiben im Rahmenplan von Mecklenburg-Vorpommern (2004)

Ein Rahmenplan ist „eine zentrale bildungspolitische Vorgabe zur Planung und Durchführung des Unterrichts, die einen *Rahmen* vorgibt, der *planvoll* zu füllen ist. Ein Rahmenplan enthält neben Verbindlichem auch Anregungen für Fakultatives . . .“ (L.I.S.A. Mecklenburg-Vorpommern, 2006, S. 37). Der Rahmenplan Grundschule Deutsch (MBS et al., 2004) ist jeweils für zwei Klassenstufen gültig. Er gilt für die erste und zweite Klassenstufe sowie für die dritte und vierte Klassenstufe. Ausgangspunkt für den Rahmenplan (MBS et al., 2004) ist der Kompetenzansatz und das damit verbundene Ziel, bei Kindern Handlungskompetenz zu entwickeln. Handlungskompetenz besteht aus Sach-, Methoden-, sozialer und personaler Kompetenz. Da die Bildungsstandards (KMK, 2005)

und der Rahmenplan zeitgleich erarbeitet wurden, konnten von Beginn an die Bildungsstandards als wesentlicher Bestandteil für den Rahmenplan einbezogen werden. Folglich werden auch im Rahmenplan vier Kompetenzbereiche bzw. Aufgabenbereiche (Begriff des Rahmenplans) unterschieden. Inhaltlich sind die Bereiche identisch, jedoch im Rahmenplan für das Lehrerhandeln ausführlicher erläutert. Den Aufgabenbereichen werden neben Inhalten Ziele zugeordnet und als Standards verbindlich festgelegt (Landesinstitut für Schule und Medien Brandenburg, 2006). Dem Aufgabenbereich „Schreiben: Texte verfassen/Rechtschreiben“ werden im Hinblick auf die Rechtschreibkompetenz für die Klassenstufe 1/2 die in Tabelle 4 dargestellten Ziele und Inhalte zugeordnet.

Tabelle 4: Ausgewählte Ziele und Inhalte aus dem Aufgabenbereich „Schreiben: Texte verfassen/Rechtschreiben“ für die Klassenstufe 1/2 (MBS et al., 2004, S. 28)

Ziele	Inhalte
Richtig schreiben	
<ul style="list-style-type: none"> - Übungswortschatz richtig schreiben - sich selbst kontrollieren und korrigieren - eigene Texte lautorientiert aufschreiben und sich an Modellwörtern orientieren - Rechtschreibmuster - Rechtschreibstrukturen berücksichtigen 	<ul style="list-style-type: none"> häufig gebrauchte Wörter, Klassenwortschatz, interessenbezogene Wörter Verfahren zur Selbstkontrolle und -korrektur Laut-Buchstabenzuordnung einfache Regeln für Groß- und Kleinschreibung, Satzschlusszeichen
Rechtschreibstrategien und Arbeitstechniken ausbilden	
<ul style="list-style-type: none"> - Wörter beim Schreiben mitsprechen - Schreibweise von Wörtern ableiten - Wörter einprägen - Wörter, Sätze, Texte sicher abschreiben - Wörter üben - Alphabet kennen und nutzen 	<ul style="list-style-type: none"> Lautfolge in Wörtern, Schreiblesen Durchgliedern von Wörtern Wortbausteine, Endungen, Wortstamm <i>Verlängern, Zurückführen auf den Wortstamm, Zerlegen, Pluralformen</i> Rechtschreibbesonderheiten, Übungsformen, Merkhilfen Abschreibhilfen, <i>Abschreiben in vier Schritten, Gliedern schwieriger Wörter</i> vielfältige Übungsformen Wörterlisten und Wörterbuch

Anmerkungen: Die fakultativen Inhalte sind kursiv dargestellt.

Der Rahmenplan enthält keinen Grundwortschatz, der jeweils am Ende einer Klassenstufe beherrscht werden soll. Es wird ein Übungswortschatz empfohlen, der aus häufig

vorkommenden, sachbezogenen und interessenbezogenen Wörtern besteht (MBS et al., 2004).

Der inhaltliche Schwerpunkt des Rahmenplans Grundschule Deutsch (MBS et al., 2004) liegt für den Lernbereich Rechtschreiben in der 1. und 2. Klassenstufe in dem Erarbeiten und Beherrschen der alphabetischen Rechtschreibstrategie. Die alphabetische Strategie ist die Grundlage für alle weiteren Rechtschreibstrategien. Nach der Sicherung der alphabetischen Strategie beginnt in der 2. Klassenstufe die Erarbeitung der orthographischen Strategie. Im Bereich Rechtschreiben stehen dann die Erarbeitung orthographischen Schreibens und die Einführung einiger Rechtschreibregeln im Mittelpunkt.

3.5.3 Rechtschreiben in Sprachbüchern

Ein Sprachbuch ist ein Buch, das im Deutschunterricht eingesetzt wird. Es enthält Texte und Aufgaben für verschiedene Lernbereiche, z. B. Grammatik, Rechtschreiben und Lesen. Der Schulbuchkatalog 2009/2010 (MBWK, 2009a) verzeichnet für die 2. Klassenstufe acht Sprachbücher:

1. „Das Auer Sprachbuch“ (Dolenc, Fiskus, Kraft, E. Röbe & H. Röbe, 2007),
2. „Das Sprachbuch 2“ (Baumann, Pristl & J. Schmidt, 2004),
3. „Jo-Jo Sprachbuch 2“ (Burgel et al., 2004),
4. „Bausteine Sprachbuch 2“ (Bauch et al., 2008),
5. „Duden Sprachbuch 2“ (Hamisch et al., 2005),
6. „Bücherwurm Sprachbuch 2“ (Ahlgrimm, Bartkowski, Karasz & Pfeiffer, 2004),
7. „Pustebume Sprachbuch 2“ (Menzel, 2008) und
8. „Sprachfreunde 2“ (Szelenko, T. Schmidt & Wessel, 2004).

„Das Auer Sprachbuch“ (Dolenc et al., 2007) wird für die erste und zweite Klassenstufe verwendet, wobei der Inhalt der 1. Klassenstufe nur die ersten 19 Seiten umfasst. Der Bereich Rechtschreiben wird parallel zu den Themen des Sprachbuches behandelt, aber am Ende des Buches nochmals zusammengefasst aufgeführt. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares graphematisches und orthographisches sowie erweitertes graphematisches Wissen. Als

elementares graphematisches Wissen werden eindeutige Phonem-Graphem-Korrespondenzen, die mehrgliedrigen Grapheme *ch*, *sch*, *qu*, *st*, *sp* und *ng* sowie die Diphthonggrapheme *ei*, *au*, *eu* und das *v* als [f] und [w] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen und die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung, die Endsilben *-en*, *-er*, *-el* und Anfangssilben *an-*, *ein-*, *ver-* sowie die Vokalisierung des *r* als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen werden in diesem Sprachbuch das Dehnungs-h und das Vokalgraphem *ie* eingeführt. Es ist zu vermuten, dass „Das Auer Sprachbuch“ (Dolenc et al., 2007) auf dem Zwei-Wege-Modell des Rechtschreibprozesses basiert, denn zum einen vermittelt es Regelwissen in Form von Merksätzen wie „Nach einem langen Selbstlaut oder Umlaut steht oft ein **h**“ (Dolenc et al., 2007, S. 127), zum anderen vermittelt es Merkwörter, die durch orthographische Besonderheiten gekennzeichnet sind. In diesem Sprachbuch werden die Arbeitstechniken des Mitsprechens, des überdeutlichen Mitsprechens, des Abschreibens, des strategiegeleiteten Einprägens und des Ableitens verwendet. Als weitere Arbeitstechnik für orthographisch korrektes Schreiben wird das Nachschlagen in der Wörterliste, die sich am Ende des Buches befindet, vermittelt.

„Das Sprachbuch 2“ (Baumann et al., 2004) behandelt den Bereich Rechtschreiben parallel zu den einzelnen Themen des Sprachbuches. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares graphematisches und orthographisches sowie erweitertes graphematisches Wissen. Als elementares graphematisches Wissen werden eindeutige Phonem-Graphem-Korrespondenzen und die mehrgliedrigen Grapheme *qu*, *st* und *sp* sowie die Diphthonggrapheme *ei*, *au*, *eu* und das *v* als [f] eingeführt. Als elementare orthographische Kompetenzen werden die Großschreibung am Satzanfang und konkreter Nomen und die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung, die Endsilben *-e*, *-en*, *-er*, *-st*, *-t*, die Anfangssilbe *ver-* und die Vokalisierung des *r* als elementare orthographische Kompetenzen vermittelt. Als erweitertes graphematisches Wissen werden in dem Sprachbuch das Dehnungs-h und das Vokalgraphem *ie* als Dehnungszeichen und die Doppelkonsonanz als Kürzezeichen eingeführt. „Das Sprachbuch 2“ (Baumann et al., 2004) basiert auf dem Zwei-Wege-Modell des Rechtschreibprozesses, denn es vermittelt Mitsprech-, Nachdenk- und Merkwörter. Mitsprechwörter mit eindeutiger Phonem-Graphem-Korrespondenz und Nachdenkwörter mit phonologischen oder morphologischen

Besonderheiten, z. B. „Verwandte Wörter suchen: **a** → **ä** [,] **au** → **äu**“ (Baumann et al., 2004, S. 84), fokussieren den indirekten Abruf von Schreibungen. Merkwörter, z. B. „Wörter mit **V**, **v** will ich mir gut merken“ (Baumann et al., 2004, S. 84), entsprechen dem direkten Abrufen von Schreibungen aus einem inneren orthographischen Lexikon. In diesem Sprachbuch werden das Mitsprechen, Silbensprechen, Abschreiben, Einprägen mit einer Wörterkartei und Ableiten als Arbeitstechniken verwendet. Als weitere Arbeitstechnik für orthographisch korrektes Schreiben wird das Nachschlagen in der Wörterliste, die sich am Ende des Buches befindet, vermittelt.

Im „Jo-Jo Sprachbuch 2“ (Burgel et al., 2004) wird der Bereich Rechtschreiben parallel zu den Themen des Sprachbuches behandelt und zusammengefasst mit weiteren Übungen im Anhang aufgeführt. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares und erweitertes graphematisches und orthographisches Wissen. Als elementares graphematisches Wissen werden die mehrgliedrigen Grapheme *ch*, *sch*, *st* und *sp* sowie die Diphthonggrapheme *ei*, *au*, *eu* eingeführt. Außerdem wird das *v* als [f] und [w] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen und die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung, die Anfangssilben *vor-*, *ab-*, *mit-*, *nach-*, *durch-*, *auf-* als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen werden in dem Sprachbuch das Dehnungs-h, das Vokalgraphem *ie* und die Doppelvokale in Nomen als Dehnungszeichen und die mehrgliedrigen Grapheme *ck* und *tz* sowie die Doppelkonsonanz als Kürzezeichen eingeführt. Das Sprachbuch führt mit dem silbenanlautenden-h, ohne dieses explizit von Wörtern mit Dehnungs-h abzugrenzen, bereits erweitertes orthographisches Wissen ein. Anhand des Aufbaus und der Übungsformen des Sprachbuches kann nicht geschlussfolgert werden, ob das „Jo-Jo Sprachbuch 2“ (Burgel et al., 2004) auf einer Theorie des Rechtschreibprozesses basiert. In dem Sprachbuch werden für orthographisch richtiges Schreiben das Abschreiben, Silbensprechen, Ableiten sowie das Nachschlagen in der Wörterliste als Arbeitstechniken vermittelt.

Das „Bausteine Sprachbuch 2“ (Bauch et al., 2008) behandelt den Bereich Rechtschreiben mit insgesamt sieben Kapiteln. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares graphematisches und

orthographisches sowie erweitertes graphematisches Wissen. Als elementares graphematisches Wissen werden die mehrgliedrigen Grapheme *ch*, *qu*, *st*, *sp*, der Buchstabe *x*, die Diphthonggrapheme *ei*, *au*, *eu* sowie das *v* als [f] und [w] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen sowie die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung, die Endsilben *-en*, *-er*, *-el* und Anfangssilben *vor-*, *ein-*, *ver-*, *zu-*, *nach-*, *weg-*, *mit-* sowie die Vokalisierung des *r* als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen werden in dem Sprachbuch das Vokalgraphem *ie*, die Doppelvokale in Nomen als Dehnungszeichen und die Doppelkonsonanz als Kürzezeichen eingeführt. Es ist zu vermuten, dass das „Bausteine Sprachbuch 2“ (Bauch et al., 2008) auf dem Zwei-Wege-Modell des Rechtschreibprozesses basiert, da es zum einen Regelwissen vermittelt, z. B. „Wenn man ein langes ‚i‘ hört, schreibt man meistens ein *ie*“ (Bauch et al., 2008, S. 64) und zum anderen durch das mehrmalige Abschreiben von Wörtern diese geübt und behalten werden sollen. Als Arbeitstechniken für orthographisch richtiges Schreiben werden das Abschreiben, Silbensprechen, Einprägen, Ableiten und die Arbeit mit der Wörterliste, die sich am Ende des Buches befindet, vermittelt.

Das „Duden Sprachbuch 2“ (Hamisch et al., 2005) behandelt den Bereich Rechtschreiben allen Themen nachgeordnet in einem eigenen Kapitel. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares und erweitertes graphematisches und orthographisches Wissen. Als elementares graphematisches Wissen werden die mehrgliedrigen Grapheme *ch*, *sch*, *qu*, *st*, *sp*, *ng*, *nk* und der Buchstabe *x* eingeführt. Des Weiteren werden die Diphthonggrapheme *ei*, *au*, *eu* und das *v* als [f] und [w] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen sowie die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung und die Anfangssilben *vor-* und *ver-* als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen werden in dem Sprachbuch das Dehnungs-h, das Vokalgraphem *ie* als Dehnungszeichen und die mehrgliedrigen Grapheme *ck* und *tz* sowie die Doppelkonsonanz als Kürzezeichen eingeführt. Das Sprachbuch führt mit dem silbenanlautenden-h und dem *ß* nach langem Vokal bereits erweitertes orthographisches Wissen ein. Das „Duden Sprachbuch 2“ (Hamisch et al., 2005) basiert auf dem Zwei-Wege-Modell des Rechtschreibprozesses. Einerseits gibt es Merksätze, die einen indirekten Abruf von

Schreibungen fokussieren, andererseits gibt es Abschreibübungen, die das direkte Abrufen von Schreibungen fokussieren sollen. In diesem Sprachbuch werden die Arbeitstechniken des Abschreibens, Einprägens und Nachschlagens in der Wörterliste, die sich am Ende des Buches befindet, für orthographisch korrektes Schreiben vermittelt.

Das „Bücherwurm Sprachbuch 2“ (Ahlgrimm et al., 2004) verknüpft den Bereich Rechtschreiben mit allen Themenbereichen. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares und erweitertes graphematisches und orthographisches Wissen. Als elementares graphematisches Wissen werden das Graphem ⟨z⟩ und die mehrgliedrigen Grapheme *ch*, *sch*, *qu*, *st*, *sp*, *ng*, *nk* eingeführt. Des Weiteren werden die Diphthonggrapheme *ei*, *au*, *eu* und das *v* als [f] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen sowie die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Außerdem werden die Auslautverhärtung und die Anfangssilben *ver-*, *an-*, *aus-*, *zu-*, *durch-*, *mit-*, *auf-*, *unter-*, *über-*, *ab-*, *nach-* als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen werden in dem Sprachbuch das Dehnungs-h und Vokalgraphem *ie* als Dehnungszeichen und die mehrgliedrigen Grapheme *ck* und *tz* sowie die Doppelkonsonanz als Kürzezeichen eingeführt. Das Sprachbuch führt mit dem *ß* nach langem Vokal bereits erweitertes orthographisches Wissen ein. Aufgrund unterschiedlicher Abschreibübungen, bei denen sich Lernwörter mit phonologischen, morphologischen oder orthographischen Besonderheiten eingeprägt werden sollen, und Merksätzen ist davon auszugehen, dass das Sprachbuch dem Zwei-Wege-Modell des Rechtschreibprozesses entspricht. Als Arbeitstechniken werden das Abschreiben, Einprägen und Nachschlagens in der Wörterliste, die sich am Ende des Buches befindet, vermittelt.

Das „Pustebblume Sprachbuch 2“ (Menzel, 2008) hat ein eigenes Kapitel für den Bereich Rechtschreiben. Das Sprachbuch entspricht nur in Teilen dem aktuellen Forschungsstand zur Rechtschreibentwicklung, denn Rechtschreibbesonderheiten werden nur lückenhaft vermittelt. Als elementares graphematisches Wissen werden die mehrgliedrigen Grapheme *qu*, *st* und *sp* und das *v* als [f] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen sowie Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung und mehrere Anfangssilben als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen

werden in dem Sprachbuch nur das Vokalgraphem *ie* als Dehnungszeichen und das mehrgliedrige Graphem *tz* sowie die Doppelkonsonanz als Kürzezeichen eingeführt. Aufgrund von Abschreibübungen und Merksätzen ist davon auszugehen, dass das Sprachbuch dem Zwei-Wege-Modell des Rechtschreibprozesses entspricht. Als Arbeitstechniken werden das Abschreiben, Einprägen und Nachschlagen in der Wörterliste, die sich wie bei allen bisherigen Sprachbüchern am Ende befindet, vermittelt.

Das Sprachbuch „Sprachfreunde 2“ (Szelenko et al., 2004) behandelt den Bereich Rechtschreiben parallel zu allen weiteren Themen. Das Sprachbuch entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung. Es vermittelt elementares und erweitertes graphematisches und orthographisches Wissen. Als elementares graphematisches Wissen werden die mehrgliedrigen Grapheme *ch*, *sch*, *st*, *sp*, *ng*, die Diphthonggrapheme *ei*, *au*, *eu* und das *v* als [f] eingeführt. Als elementares orthographisches Wissen werden die Großschreibung am Satzanfang und konkreter Nomen sowie die Umlautgrapheme *ä*, *ö*, *ü*, *äu* eingeführt. Des Weiteren werden die Auslautverhärtung und die Anfangssilben *ver-*, *an-*, *aus-*, *zu-*, *mit-*, *um-*, *ein-*, *be-*, *nach-*, *vor-*, *weg-* als elementares orthographisches Wissen vermittelt. Als erweitertes graphematisches Wissen werden in dem Sprachbuch das Dehnungs-h, das Vokalgraphem *ie* und die Doppelvokale als Dehnungszeichen und die mehrgliedrigen Grapheme *ck* und *tz* sowie die Doppelkonsonanz als Kürzezeichen eingeführt. Das Sprachbuch führt mit dem *ß* nach langem Vokal und dem silbenanlautenden-h, ohne dieses explizit von Wörtern mit Dehnungs-h abzugrenzen, erweitertes orthographisches Wissen ein. Das Sprachbuch führt Regelwissen über kurze Merksätze ein und fordert das Abschreiben und Einprägen von Lernwörtern mit Rechtschreibbesonderheiten. Deshalb ist anzunehmen, dass das Sprachbuch auf dem Zwei-Wege-Modell des Rechtschreibprozesses basiert. Als Arbeitstechniken werden das Abschreiben, das Einprägen mit Hilfe einer Wörterkartei, das Ableiten und das Nachschlagen in der Wörterliste, die sich am Ende des Buches befindet, vermittelt.

Die Sprachbücher aus dem Schulbuchkatalog 2009/2010 für das Unterrichtsfach Deutsch der Klassenstufe 2 in Grundschulen in Mecklenburg-Vorpommern (MBWK, 2009a) entsprechen bis auf das „Pustebblume Sprachbuch 2“ (Menzel, 2008) dem aktuellen Forschungsstand zur Rechtschreibentwicklung und zum Rechtschreibprozess. Der inhaltliche Schwerpunkt aller Sprachbücher liegt auf dem elementaren graphematischen

und orthographischen Können.

3.5.4 Zusammenfassung und Schlussfolgerungen

Sowohl die Anforderungen und Inhalte des Aufgabenbereichs Rechtschreiben in den Bildungsstandards (KMK, 2005) und im Rahmenplan Grundschule Deutsch (MBS et al., 2004) als auch die Sprachbücher (MBWK, 2009) orientieren sich an den aktuellen Theorien des Rechtschreiberwerbs. Trotz kleinerer Abweichungen in den zu vermittelnden Inhalten der Sprachbücher legen sie für die zweite Klassenstufe der Grundschule fest, dass die alphabetische und orthographische Strategie im Unterrichtsfach Deutsch erworben werden sollte. Tabelle 5 gibt einen Überblick über das zu erwerbende Rechtschreibkönnen in der 2. Klassenstufe.

Tabelle 5: Anforderungen und Inhalte des Lernbereichs Rechtschreiben in der 2. Klassenstufe

	Teilfähigkeit der Rechtschreibleistung	Anforderungen und Inhalte des Lernbereichs Rechtschreiben in der 2. Klassenstufe
Elementare Stufe	Graphematisches Können	<ul style="list-style-type: none"> - Eindeutige Phonem-Graphem-Korrespondenzen - Schreibung des [ʃ] als <i>sch</i>, <i>st</i>, <i>sp</i> - Schreibung des [ç] und [x] als <i>ch</i> - Schreibung von <i>ei</i>, <i>au</i>, <i>eu</i> - Differenzierung zwischen <i>ng</i> und <i>nk</i> - Schreibung des Graphems <i>qu</i>
	Orthographisches Können	<ul style="list-style-type: none"> - Großschreibung am Satzanfang - Großschreibung konkreter Nomen - Kleinschreibung von Verben und Adjektiven - Schreibung des [f] und [w] als <i>v</i> - Schreibung der Anfangssilben <i>ver-</i>, <i>vor-</i>, <i>auf-</i>, <i>an-</i>, <i>zu-</i> - Schreibung der Endsilben <i>-en</i>, <i>-er</i>, <i>-el</i> - Morphemkonstanz: <i>ä</i>, <i>ö</i>, <i>ü</i>, <i>äu</i> - Auslautverhärtung - Satzschlusszeichen: <i>.</i>, <i>?</i>, <i>!</i>
Erweiterte Stufe	Graphematisches Können	<ul style="list-style-type: none"> - Dehnungs-h als Dehnungszeichen - <i>ie</i> als Dehnungszeichen - Doppelkonsonanz als Kürzezeichen - <i>ck</i> als Kürzezeichen - <i>tz</i> als Kürzezeichen
		- Arbeitstechniken: Abschreiben, Nachschlagen, Einprägen, Ableiten, Selbstkontrolle, -korrektur

Aus den dargestellten Anforderungen und Inhalten des Lernbereichs Rechtschreiben und der Frage, welche genannten Teilleistungen messbar sind, ergibt sich für die Entwicklung der FE-RS 2 Folgendes: Die Überprüfung elementaren graphematischen und orthographischen sowie erweiterten graphematischen Könnens ist über ein Diktat einzelner Wörter und Sätze oder Lückentexte möglich. Dabei kann auch der genannte Aspekt der Zeichensetzung berücksichtigt werden. Die Überprüfung der Arbeitstechniken ist nur bedingt möglich. Die Fähigkeit, in Wörterbüchern nachzuschlagen, könnte über das alphabetische Sortieren von Wörtern überprüft werden. Ob Schreibungen abgeleitet oder eingeprägt werden, ist, wie bereits erwähnt, nicht überprüfbar. Die Überprüfung von Selbstkontrolle und -korrektur würde eher Rechtschreibwissen als Rechtschreibkönnen überprüfen (Böhme & Bremerich-Vos, 2009). Schlussfolgernd soll das zu entwickelnde Verfahren zur Lernverlaufsdokumentation überprüfen, inwieweit Kinder der 2. Klassenstufe die dargestellten Anforderungen und Inhalte des elementaren graphematischen und orthographischen sowie des erweiterten graphematischen Rechtschreibkönnens beherrschen. Die Überprüfung der Arbeitstechniken soll nicht erfolgen. Die Auswahl der Testitems, das heißt das Wortmaterial, soll auf den Wörterlisten der Sprachbücher beruhen. Die Testitems sollen auf den orthographischen Besonderheiten elementaren graphematischen und orthographischen sowie erweiterten graphematischen Könnens beruhen. Das bedeutet, dass den Schülern zu jedem Messzeitpunkt Wortmaterial unterschiedlichen Schwierigkeitsgrades, von leicht bis schwer, von elementar graphematisch bis erweitert graphematisch, vorgelegt werden soll.

4 Methodik der Untersuchung

Das Hauptziel der Studie ist die theoriegeleitete Entwicklung und Evaluation eines veränderungssensitiven Verfahrens zur Lernverlaufsdokumentation für den Lernbereich Rechtschreiben der 2. Klassenstufe; der FE-RS 2. Dieses Untersuchungsziel wurde in den vorangegangenen Kapiteln konkretisiert, indem Aussagen zur Konstruktion der FE-RS 2 aus dem Forschungsstand zu ATI und RTI sowie zum Rechtschreiberwerb abgeleitet wurden. In diesem Kapitel wird die Methodik der Untersuchung dargelegt. Dazu zählen, dass der Untersuchungsplan, die in der Studie verwendeten Untersuchungsinstrumente, die Stichprobengewinnung und die Untersuchungsdurchführung beschrieben und die abgeleiteten Forschungsfragen und Hypothesen sowie die verwendeten statistischen Analyseverfahren dargelegt werden.

4.1 Untersuchungsplan

Die vorliegende Untersuchung ist eine experimentelle Felduntersuchung, bei der das natürliche Umfeld, die Schule und der Unterricht, nicht beeinflusst wird. Der Untersuchung liegt ein einfaktorieller, univariater Zweistichprobenversuchsplan mit viermaliger Messwiederholung im Cross-Over-Design zugrunde. Ein einfaktorieller Versuchsplan untersucht den Einfluss *einer* unabhängigen Variable, hier der Faktor Zeit, auf abhängige Variablen (Bortz & Schuster, 2010). Ein univariater Versuchsplan untersucht den Einfluss unabhängiger Variablen auf *eine* abhängige Variable, hier die Entwicklung der Rechtschreibleistung (Bortz & Döring, 2006). Mit einem einfaktoriellen Versuchsplan mit Messwiederholung lassen sich Lernfortschritte *n*-fach wiederholt untersuchen, indem *n* Messungen einer Stichprobe zu verschiedenen Zeitpunkten erhoben werden (Zendler, Spannagel & Vogel, 2008). So werden in der vorliegenden Untersuchung von der Stichprobe nicht nur eine, sondern vier Messungen im Abstand von 10 Wochen erhoben. Bei einem einfaktoriellen Zweistichprobenplan wird die Gesamtstichprobe in zwei möglichst gleich große Teilstichproben randomisiert aufgeteilt, um die unterschiedlichen Ausgangssituationen beider Teilstichproben zu minimieren (Bortz & Döring, 2006). So wird auch in der vorliegenden Untersuchung die Gesamtstichprobe in zwei gleichgroße Gruppen von Schulklassen randomisiert aufgeteilt. Das Cross-Over-Design wird in der vorliegenden Untersuchung gewählt, da dieselbe Stichprobe mehrfach

untersucht wird, um Lernverläufe erfassen zu können (Bortz & Schuster, 2010) und die FE-RS 2 in den zwei Parallelförm A und B vorliegt. Für die Untersuchung resultiert daraus ein 2 x 2-Quadrat mit Messwiederholung, das heißt die erste Gruppe von Schulklassen führt zuerst die FE-RS 2 in der Form A und danach in der Form B durch, für die zweite Gruppe von Schulklassen ist die Reihenfolge umgekehrt. Da es sich um einen Versuchsplan mit Messwiederholung handelt, ergibt sich das der Untersuchung zugrunde liegende Untersuchungsdesign, welches in Tabelle 6 dargestellt ist.

Tabelle 6: Untersuchungsdesign

	1. MZP	2. MZP	3. MZP	4. MZP
1. Gruppe von Schulklassen	A	B	A	B
2. Gruppe von Schulklassen	B	A	B	A

Anmerkungen: MZP = Messzeitpunkt, A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Die Untersuchung wurde in 2. Klassenstufen von Grundschulen und Regionalen Schulen mit Grundschulteil im Schulamtsbezirk Neubrandenburg in Mecklenburg-Vorpommern von November 2009 bis Juni 2010 durchgeführt. Die Untersuchung umfasst, im Abstand von jeweils 10 Wochen, 4 Messzeitpunkte. Zum 1. Messzeitpunkt wurden die Rechtschreib- und Leseleistungen der Schüler mittels der FE-RS 2, der am Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation der Universität Rostock entwickelten „Formativen Erfassung der Lesefertigkeit im 2. Schuljahr“ (FE-L 2) (Kuhlmann & Hartke, 2009b) und des DERET 1-2+ (Stock & Schneider, 2008a) ermittelt. Dabei führte die erste Gruppe der untersuchten Schulklassen entsprechend des Untersuchungsdesigns die FE-RS 2 in der Form A und die zweite Gruppe die FE-RS 2 in der Form B durch. In einer kleinen Gruppe von sechs Klassen wurde des Weiteren überprüft, ob es sich bei den Formen A und B der FE-RS 2 um Parallelförm handelt, indem innerhalb von einer Woche beide Formen durchgeführt wurden. Darüber hinaus wurden die Rechenleistungen der Schüler mittels des am Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation der Universität Rostock entwickelten Verfahrens „Formative Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr“ (FE-AF 2) (Kuhlmann & Hartke, 2009a) und des Deutschen Mathematiktests für erste Klassen (DEMAT 1+) (Krajewski, Küspert & Schneider, 2002) erfasst. Die Entwicklung und Evaluation der FE-L 2 und FE-AF 2 sind jedoch nicht Teil der Dissertation, sondern wurden nur im Rahmen der vorliegenden Untersuchung durchgeführt. Zum 2., 3. und 4.

Messzeitpunkt wurden die FE-RS 2, die FE-L 2 und die FE-AF 2 entsprechend des Untersuchungsdesigns eingesetzt. Zum 4. Messzeitpunkt wurden außerdem die Rechtschreibleistungen der Schüler mit dem DERET 1-2+ und die Rechenleistungen mit dem Deutschen Mathematiktest für zweite Klassen (DEMAT 2+) (Krajewski, Liehm & Schneider, 2004) erfasst. Zusätzlich wurde zu diesem Messzeitpunkt in einer kleinen Gruppe von fünf Klassen erneut überprüft, ob es sich bei den Formen A und B der FE-RS 2 um Parallelförmchen handelt, indem innerhalb von einer Woche beide Formen durchgeführt wurden. Des Weiteren wurden zu diesem Messzeitpunkt ein Lehrerfragebogen zur Einschätzung der FE-RS 2 und der FE-AF 2 sowie Lehrerfragebögen zur Einschätzung der Rechtschreibleistung und der arithmetischen Leistung der Schüler eingesetzt. Im Rahmen der vorliegenden Arbeit werden lediglich die Lehrerangaben zur FE-RS 2 ausgewertet, denn die Auswertung des Lehrerfragebogens zur Einschätzung der arithmetischen Leistung der Schüler ist nicht Teil der Dissertation. Abbildung 15 gibt die zeitliche und inhaltliche Struktur der Untersuchung zusammenfassend wieder.

1. MZP 10. Schulwoche (02.11.-06.11.2009)	2. MZP 20. Schulwoche (11.01.-15.01.2010)	3. MZP 30. Schulwoche (22.03.-26.03.2010)	4. MZP 40. Schulwoche (31.05.-04.06.2010)
FE-RS 2, Form A/B FE-L 2 DERET 1-2+, Klasse 1 Form A FE-AF 2 DEMAT 1+, Form A	FE-RS 2, Form A/B FE-L 2 FE-AF 2	FE-RS 2, Form A/B FE-L 2 FE-AF 2	FE-RS 2, Form A/B FE-L 2 DERET 1-2+, Klasse 2 Form A FE-AF 2 DEMAT 2+, Form A Lehrerfragebogen zur Einschätzung der FE-RS 2 und der FE-AF 2 Lehrerfragebögen zur Einschätzung der Rechtschreib- leistung sowie der arithmetischen Leistung der Schüler

Abbildung 15: Untersuchungsplan

Anmerkungen: MZP = Messzeitpunkt, FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A/B, FE-L 2 = Formative Erfassung der Lesefertigkeit im 2. Schuljahr (Kuhlmann & Hartke, 2009b), DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), FE-AF 2 = Formative Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr (Kuhlmann & Hartke, 2009a), DEMAT 1+ = Deutscher Mathematiktest für erste Klassen (Krajewski et al., 2002), DEMAT 2+ = Deutscher Mathematiktest für zweite Klassen (Krajewski et al., 2004)

4.2 Untersuchungsinstrumente

Nachfolgend werden die in der Untersuchung verwendeten und die Dissertation betreffenden Untersuchungsinstrumente beschrieben. Es wird der Deutsche Rechtschreibtest für das erste und zweite Schuljahr von Stock und Schneider (2008a), der zur Prüfung der externen Validität der FE-RS 2 herangezogen wurde, beschrieben. Zudem wird der Konstruktionsprozess der FE-RS 2, des Lehrerfragebogens zur Einschätzung der diagnostischen Verfahren zur Lernverlaufsdokumentation und des Lehrerfragebogens zur Einschätzung der Rechtschreibleistung der Schüler dargelegt.

4.2.1 Deutscher Rechtschreibtest für das erste und zweite Schuljahr von Stock und Schneider (2008a)

Der DERET 1-2+ von Stock und Schneider (2008a) dient der Erfassung und Überprüfung der Rechtschreibleistung von Kindern am Ende der 1. und 2. Klassenstufe sowie zu Beginn der 2. und 3. Klassenstufe. Das Testverfahren kann als Gruppen- und Einzeltest durchgeführt werden. Die Durchführung beansprucht 30 bis 45 Minuten. Für einzelne Items bestehen keine Zeiteinschränkungen. Die theoretischen Grundlagen für die Konzeption des DERET 1-2+ (Stock & Schneider, 2008a) bilden das Stufenmodell der Schriftsprachentwicklung von Frith (1985) und das Modell der zweifachen Zugangswege beim Rechtschreiben nach Diktat von Treiman (1991). Zudem werden linguistische Aspekte der Schriftsprache und orthographische Prinzipien berücksichtigt. Für die Konstruktion des DERET 1-2+ wurden zudem die Lehrplaninhalte aller Bundesländer und die gängigsten Schulbücher für das Unterrichtsfach Deutsch herangezogen (Stock & Schneider, 2008a, 2008b).

Der DERET 1-2+ (Stock & Schneider, 2008a) prüft die Rechtschreibleistungen von Kindern durch Fließ- und Lückentexte mit Illustrationen. Es liegen zwei Parallelformen, Form A und B, vor. Die Fließtexte bestehen für die zweite Klassenstufe aus 52 Wörtern. Die Lückentexte bestehen für die zweite Klassenstufe aus 12 Lückentextwörtern. Für das Testverfahren wurden die häufigsten Wörter aus Schulbuchwörterlisten und aus den in Lehrplänen vorhandenen Grundwortschätzen berücksichtigt und einzeln ausgewählt. Zusätzlich wurden Wörter ausgewählt, die möglichst in keiner Liste aufgeführt waren, um Transferleistungen zu überprüfen (Stock & Schneider, 2008a, 2008b).

Die Auswertung erfolgt über eine quantitative und qualitative Fehleranalyse. Die Fließtexte dienen der quantitativen, die Lückentexte der qualitativen Einschätzung der Rechtschreibleistung. Bei der quantitativen Fehleranalyse werden im Fließtext alle Falschschreibungen markiert. Als Falschschreibungen gelten ausgelassene Wörter, Oberzeichenfehler wie fehlende i-Punkte und Umlaut-Markierungen, das Fehlen des Trennungsstrichs, nicht erkennbare Wortschreibungen, verschieden dargebotene Schreibungen für ein Wort und die falsche Groß- und Kleinschreibung. Oberzeichenfehler, Fehler bei der Großschreibung am Satzanfang und beim Satzschlusszeichen sowie Übergeneralisierungen in der Zeichensetzung werden extra

erfasst. Bei der qualitativen Fehleranalyse werden im Lückentext alle Falschschreibungen markiert. Als Falschschreibungen gelten Wortauslassungen, falsche Schreibung von Buchstaben, Getrennt- und Zusammenschreibung, Groß- und Kleinschreibungsfehler, Unterscheidung von w- und f-Lauten, Dehnungs- und Kürzungsfehler, Fehler bei speziellen Vokalschreibweisen, Fehler bei Zisch- und Rachenlauten, fehlerhafte Wortendungen und sonstige nicht-lautgetreue Rechtschreibfehler. Für die Auswertungen liegen Auswertungsbögen vor (Stock & Schneider, 2008a, 2008b).

Die Durchführungsanleitung und die Vorgaben zur Auswertung und Interpretation des DERET 1-2+ (Stock & Schneider, 2008a) gewährleisten die Objektivität des Rechtschreibtests. Die Reliabilitätsbestimmung für die Fließtexte des DERET 1-2+ erfolgt über die Berechnung der internen Konsistenz (Cronbachs alpha), der Halbierungsreliabilität nach Spearman-Brown sowie der Retest- und Paralleltestreliabilität. Die interne Konsistenz und Halbierungsreliabilität erreichen für die Rechtschreibfehler beider Testformen und Klassenstufen Werte von $r = .89$ bis $r = .92$. Diese hohen Werte bestätigen den inneren Zusammenhang zwischen den Items und die Homogenität. Für die Bestimmung der Retestreliabilität wurde die jeweilige Testform im Abstand von sechs Wochen ein zweites Mal vorgelegt. Für die Bestimmung der Paralleltestreliabilität wurde die jeweils andere Testform im Abstand von vier Wochen bearbeitet. Die Retest- und Paralleltestreliabilität liegen für beide Testformen und Klassenstufen zwischen $r = .82$ und $r = .93$. Diese mittleren bis hohen Werte bestätigen die Testzuverlässigkeit und die Vergleichbarkeit beider Testformen. Insgesamt kann die Rechtschreibleistung von Kindern der 1. und 2. Klassenstufe mit dem DERET 1-2+ mit einer hohen Genauigkeit gemessen werden. Die inhaltliche Validität des Verfahrens ist durch die Analyse der Lehrplananforderungen aller Bundesländer und der Schulbuchinhalte gewährleistet. Die kriterienbezogene Validität wurde mit dem Diagnostischen Rechtschreibtest für 1. Klassen (DRT 1) (R. Müller, 1990) und DRT 2 (R. Müller, 1997), der Würzburger Leiseleseprobe (WLLP) von Küspert und Schneider (1998), dem Grundintelligenztest Skala 1 (CFT 1) von Cattell, Weiß und Osterland (1997) sowie dem Lehrerurteil überprüft. Die Korrelationen zwischen den beiden Testformen des DERET 1-2+ und den verschiedenen Verfahren liegen zwischen $r = .52$ und $r = .82$. Diese mittleren bis hohen Werte bestätigen die valide Erfassung der Rechtschreibleistung durch den DERET 1-2+ (Stock & Schneider, 2008a, 2008b).

Die bundesweite Normierung fand von Ende Mai bis Ende Juli sowie im September und Oktober 2003 an 2 095 Kindern der 1. und 2 009 Kindern der 2. Klassenstufe statt. Auf Basis dieser Daten wurden Prozentränge für die quantitative und qualitative Analyse, getrennt nach Einzel- und Klassennorm sowie nach Testform und Klassenstufe, ermittelt (Stock & Schneider, 2008a).

4.2.2 Entwicklung und Erprobung des Verfahrens „Formative Erfassung der Rechtschreibleistung im 2. Schuljahr (FE-RS 2)“

Mit der Entwicklung der FE-RS 2 wird eine rechtschriftliche Diagnostik auf der elementar graphematischen (alphabetische Strategie), der elementar orthographischen (beginnend orthographische Strategie) und der erweitert graphematischen (zunehmend orthographische Strategie) Stufe angestrebt, da diese Strategien die zweite Klassenstufe dominieren (Naumann, 2008; Scheerer-Neumann, 2003, 2006). Die FE-RS 2 soll ein Gruppentest sein, der aber auch in einer Einzelsituation durchgeführt werden kann. Die FE-RS 2 soll ein zeitökonomisches Verfahren sein, das in einer Unterrichtsstunde durchzuführen ist und somit einschließlich der Instruktionen zu Beginn der 2. Klassenstufe maximal 40 Minuten und im weiteren Verlauf des Schuljahres deutlich weniger Zeit in Anspruch nimmt. Gleichzeitig sollte die Anzahl der Testwörter und Rechtschreibbesonderheiten relativ hoch sein, um die Bandbreite der curricularen und theoretischen Vorgaben abzudecken. Die Auswahl der Testwörter sollte nicht beliebig sein. Deshalb wurden für die FE-RS 2 die Testwörter für zwei Parallelversionen vor dem Hintergrund linguistischer und entwicklungstheoretischer Aspekte aus den Wörterlisten der Sprachbücher (MBWK, 2009a) ausgewählt. Aufgrund dieser Überlegungen wurden aus den über 2 000 Wörtern der acht Sprachbücher die Wörter ausgewählt, die einem theoretisch fundierten Kategoriensystem, das nach Silbenstruktur und Rechtschreibbesonderheiten aufgebaut war, entsprachen. Dem Kategoriensystem entsprechend wurden 57 geeignete Wörter ausgewählt und fünf weitere Wörter (für die Form A: Zahn, Speise und für die Form B: kahl, Tomaten, Beule) ergänzt. Diese insgesamt 62 Wörter entsprechen den curricularen Vorgaben und entwicklungstheoretischen Annahmen und weisen sogenannte *Lupenstellen* für die oben genannten Strategien auf. Das Konzept der Lupenstellen entstammt der HSP von May (2002). Lupenstellen markieren orthographische Besonderheiten in Wörtern und „die Werte für die Lupenstellen zeigen den Grad der Beherrschung der

Rechtschreibstrategien . . .“ (May, 2002, S. 12). Entsprechend des dieser Arbeit zugrunde liegenden Entwicklungsmodells graphematischen und orthographischen Rechtschreibkönnens soll die FE-RS 2 zwischen Lupenstellen der elementar graphematischen und orthographischen sowie der erweitert graphematischen und orthographischen Stufe differenzieren. Für die Konstruktion der FE-RS 2 wurden anschließend die 62 Wörter entsprechend geschätzter gleicher Schwierigkeitsgrade auf beide Formen A und B verteilt und nach der Schwierigkeit der Wörter, von leicht nach schwer, sortiert, sodass zwei augenscheinlich parallele Formen der FE-RS 2 mit je 31 Wörtern entstanden. Tabelle 7 gibt das Wortmaterial der FE-RS 2, geordnet nach Form A und B, und die Lupenstellen wieder.

Tabelle 7: Wortmaterial und Lupenstellen der FE-RS 2

Nr.	FE-RS 2 Form A	FE-RS 2 Form B	Silbenstruktur	ELG		ELO		ERG		ERO	
				A	B	A	B	A	B	A	B
1	Oma	Opa	V.KV	⟨oma⟩	⟨opa⟩	⟨O⟩	⟨O⟩				
2	Ofen	Esel	V.KVK	⟨ofen⟩	⟨esel⟩	⟨O⟩	⟨E⟩				
3	Nase	Name	KV.KV	⟨nase⟩	⟨name⟩	⟨N⟩	⟨N⟩				
4	Maus	Baum	KVK	⟨au⟩	⟨au⟩	⟨M⟩	⟨B⟩				
5	Meise	Seife	KV.KV	⟨ei⟩	⟨ei⟩	⟨M⟩	⟨S⟩				
6	reden	legen	KV.KVK	⟨reden⟩	⟨legen⟩						
7	Telefon	Tomaten	KV.KV.KVK	⟨telefon⟩	⟨tomaten⟩	⟨T⟩	⟨T⟩				
8	Krone	Blume	KKV.KV	⟨krone⟩	⟨blume⟩	⟨K⟩	⟨B⟩				
9	fragen	tragen	KKV.KVK	⟨fr⟩	⟨tr⟩	⟨en⟩	⟨en⟩				
10	schlafen	schlagen	KKV.KVK	⟨schl⟩	⟨schl⟩	⟨en⟩	⟨en⟩				
11	flach	frech	KKVK	⟨fl⟩ ⟨ch⟩	⟨fr⟩ ⟨ch⟩						
12	hart	kalt	KVKK	⟨hart⟩	⟨kalt⟩						
13	Wolke	Hefte	KVK.KV	⟨lk⟩	⟨ft⟩	⟨W⟩	⟨H⟩				
14	helfen	halten	KVK.KVK	⟨lf⟩	⟨lt⟩	⟨en⟩	⟨en⟩				
15	Stern	Sturm	KKVKK	⟨st⟩	⟨st⟩	⟨S⟩	⟨S⟩				
16	Speise	Spule	KKV.KV	⟨sp⟩	⟨sp⟩	⟨S⟩	⟨S⟩				
17	Qualm	Quark	KKVKK	⟨qu⟩	⟨qu⟩	⟨Q⟩	⟨Q⟩				
18	Leute	Beule	KV.KV	⟨eu⟩	⟨eu⟩	⟨L⟩	⟨B⟩				
19	singen	fangen	KVK.KVK	⟨ng⟩	⟨ng⟩	⟨en⟩	⟨en⟩				
20	Bank	Onkel	KVKK/ VK.KVK	⟨nk⟩	⟨nk⟩	⟨B⟩	⟨O⟩ ⟨el⟩				
21	Sieb	Lied	KVK			⟨S⟩ ⟨b⟩	⟨L⟩ ⟨d⟩	⟨ie⟩	⟨ie⟩		
22	zahn	kahl	KVK	⟨z⟩				⟨ah⟩	⟨ah⟩		
23	nett	nass	KVK					⟨tt⟩	⟨ss⟩		
24	lieb	tief	KVK			⟨b⟩		⟨ie⟩	⟨ie⟩		
25	Katze	Tatze	KVK.KV			⟨K⟩	⟨T⟩	⟨tz⟩	⟨tz⟩		
26	backen	packen	KVK.KVK			⟨en⟩	⟨en⟩	⟨ck⟩	⟨ck⟩		
27	Bäume	Mäuse	KV.KV			⟨B⟩ ⟨äu⟩	⟨M⟩ ⟨äu⟩				
28	vor	von	KVK			⟨v⟩	⟨v⟩				
29	Hüte	Züge	KV.KV		⟨z⟩	⟨H⟩ ⟨ü⟩	⟨Z⟩ ⟨ü⟩				
30	Vase	Vampir	KV.KV/ KVK.KVK			⟨V⟩ ⟨v⟩	⟨V⟩ ⟨v⟩				
31	Frühlüher	Fahrräder	KKV.KKV.KVK/ KVK.KV.KVK	⟨fr⟩ ⟨bl⟩		⟨F⟩ ⟨er⟩	⟨F⟩ ⟨ä⟩ ⟨er⟩	⟨üh⟩	⟨ah⟩	⟨üh⟩	⟨rr⟩

Anmerkungen:

Nr. = Nummer, FE-RS 2 Form A, A/Form B, B= Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A/B, K = Konsonant, V = Vokal, ELG = Lupenstelle der elementar graphematischen Stufe, ELO = Lupenstelle der elementar orthographischen Stufe, ERG = Lupenstelle der erweitert graphematischen Stufe, ERO = Lupenstelle der erweitert orthographischen Stufe

Die Formen A und B der FE-RS 2 sind so aufgebaut, dass das Wortmaterial von oben nach unten bezüglich der Silbenstruktur komplexer und bezüglich der Rechtschreibbesonderheiten schwieriger wird. Die Wörter Nr. 1-7 haben eine einfache Silbenstruktur ohne Konsonantenhäufung. Bis auf die Wörter Nr. 5, *Meise* und *Seife*, und Nr. 6, *reden* und *legen*, können diese Wörter lautgetreu geschrieben werden. Die Wörter Nr. 8-17 und Nr. 19-20 sind in ihrer Silbenstruktur mit initialer, medialer und/oder finaler Konsonantenhäufung von mittlerer Schwierigkeit. Das Wort Nr. 18, *Leute*, ist silbisch einfach. Aufgrund der Konsonantenhäufungen, der mehrgliedrigen Grapheme, des Diphthongs *eu* und der Differenzierung zwischen *ng* und *nk* sind die Wörter Nr. 8-20 in ihrer Rechtschreibung mittelschwierig. Die Wörter Nr. 21-24 haben eine einfache, die Wörter Nr. 25-26 eine mittelschwierige Silbenstruktur. Alle diese Wörter, Nr. 21-26, sind in ihrer Rechtschreibung aufgrund der orthographischen Kennzeichnung der Vokalquantität sehr schwierig. Die Wörter Nr. 27-30 haben eine einfache Silbenstruktur, das Wort *Vampir* ist jedoch von mittlerer Schwierigkeit. Diese Wörter sind bezüglich ihrer Rechtschreibbesonderheiten zwischen mittelschwierig und schwierig einzuordnen. Das Wort Nr. 31 ist aufgrund seiner Dreisilbigkeit und Konsonantenhäufigkeit silbisch sehr komplex. In Bezug auf seine Rechtschreibung ist das Wort sehr schwierig.

Ebenso plausibel wie das Wortmaterial soll das Auswertungssystem der FE-RS 2 sein. Die FE-RS 2 soll ein objektives, valides und reliables Globalmaß und eine qualitative Analyse für die Rechtschreibleistung der Schüler in der 2. Klassenstufe liefern. Für die Entwicklung der FE-RS 2 wurde deshalb ein curriculumorientiertes, entwicklungstheoretisch und linguistisch fundiertes Auswertungssystem zugrunde gelegt, das auf den Rechtschreibstrategien und den Lupenstellen basiert. Für die Auswertung und Analyse der geschriebenen Wörter gilt, dass die orthographisch richtig geschriebenen Wörter und die korrekten Schreibungen der Lupenstellen ausgezählt werden. In dem Auswertungssystem werden fünf Skalen voneinander unterschieden, eine Skala für den orthographischen Gesamtwert und vier Subskalen. Die Skala für den orthographischen Gesamtwert dient der quantitativen Auswertung. Die Subskalen umfassen die Lupenstellen und dienen der qualitativen Auswertung. Die fünf Skalen werden nachfolgend erläutert:

1. Skala „*Orthographischer Gesamtwert (OG)*“: Hier geht es um die orthographisch korrekte Schreibung von Wörtern, wobei fehlende i-Punkte nicht, wohl aber bei Diphthongen, als Fehler gelten.
2. Subskala „*Elementar Graphematisches Können (ELG)*“: Hier geht es vor allem um

alphabetische Schreibungen von Wörtern sowie um besondere Konsonantengrapheme wie z. B. *qu* oder *sch* und Konsonantenhäufungen. Fehlende i-Punkte, spiegelverkehrte Grapheme, ausgenommen *b* und *d*, und verkehrte Grapheme wie ⟨shc⟩ für *sch* und ⟨hc⟩ für *ch* gelten nicht als Fehler.

3. Subskala „*Elementar Orthographisches Können* (ELO)“: Hier geht es um die Groß- und Kleinschreibung, die Endsilben *-en* und *-er* sowie um vokalische Ableitungen, Auslautverhärtungen und die Schreibung des [f] und [w] als *v*.
4. Subskala „*Erweitert Graphematisches Können* (ERG)“: Hier geht es um die Vokalquantität durch die Kürze- und Dehnungszeichen, die Doppelkonsonanz, *tz*, *ck*, *ie* und Dehnungs-*h*.
5. Subskala „*Erweitert Orthographisches Können* (ERO)“: Hier geht es um das silbenanlautende-⟨h⟩ und das Prinzip der Morphemkonstanz.

Für eine objektive Durchführung der FE-RS 2 liegt ein Lehrerheft mit standardisierten Durchführungshinweisen vor. Es enthält Angaben zu den Erprobungszeiträumen, der Durchführungsdauer, dem Testmaterial, zu den Durchführungshinweisen sowie zu den wörtlichen Instruktionen für die Durchführung. Das Lehrerheft befindet sich im Anhang A. Für die Schüler wurde ein Heft in DIN A5 erstellt, in welches sie die diktierten Wörter schreiben. Das Schülerheft für die FE-RS 2 besteht aus einem Deckblatt und drei nummerierten Seiten mit grau hinterlegter Lineatur. Auf dem Deckblatt ist Platz für Personenangaben und für ein grau hinterlegtes Beispieltäfelchen. Im Inneren des Heftes sollen die Schüler die 31 Wörter untereinander schreiben, die Lineatur ist durchnummeriert. Das Schülerheft befindet sich im Anhang B.

Die Erprobungsstudie der FE-RS 2 fand in der 4. Schulwoche im Schuljahr 2009/2010 in Mecklenburg-Vorpommern statt. Insgesamt nahmen 88 Schüler von jeweils zwei 2. und 3. Klassen Rostocker Grundschulen teil. Der Zeitpunkt der Teilnahme der 2. Klassen entsprach ungefähr dem 1. Messzeitpunkt und die der 3. Klassen ungefähr dem 4. Messzeitpunkt der Untersuchung. Tabelle 8 gibt die Erprobungstichprobe der FE-RS 2 wieder.

Tabelle 8: Erprobungsstichprobe der FE-RS 2

GS Gehlsdorf, Klasse 2b N	21	GS Gehlsdorf, Klasse 3 N	22
GS Toitenwinkel, Klasse 2a N	22	GS Toitenwinkel, Klasse 3a N	23
Gesamt N	43		45

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, GS = Grundschule, N = Anzahl

Die Erprobung diente vorwiegend dazu, das Verständnis der Testanweisungen und die Schwierigkeit der Wörter in Hinblick auf unterschiedliche Schwierigkeitsgrade zu überprüfen. Des Weiteren sollte überprüft werden, ob eine Unterrichtsstunde von 45 Minuten für die Durchführung der FE-RS 2 ausreichend ist und ob die Schüler mit dem Schreiben von 31 Wörtern überfordert sind. Während der Erprobung wurde das Verhalten der Schüler beobachtet und teilweise protokolliert. Anschließend wurden, getrennt nach Klassenstufe sowie Form A und B der FE-RS 2, die Häufigkeitsverteilungen erstellt und die Mittelwerte und Standardabweichungen sowie das Minimum und Maximum berechnet.

Die Häufigkeitsverteilung für den orthographischen Gesamtwert der Form A ist in Abbildung 16 dargestellt.

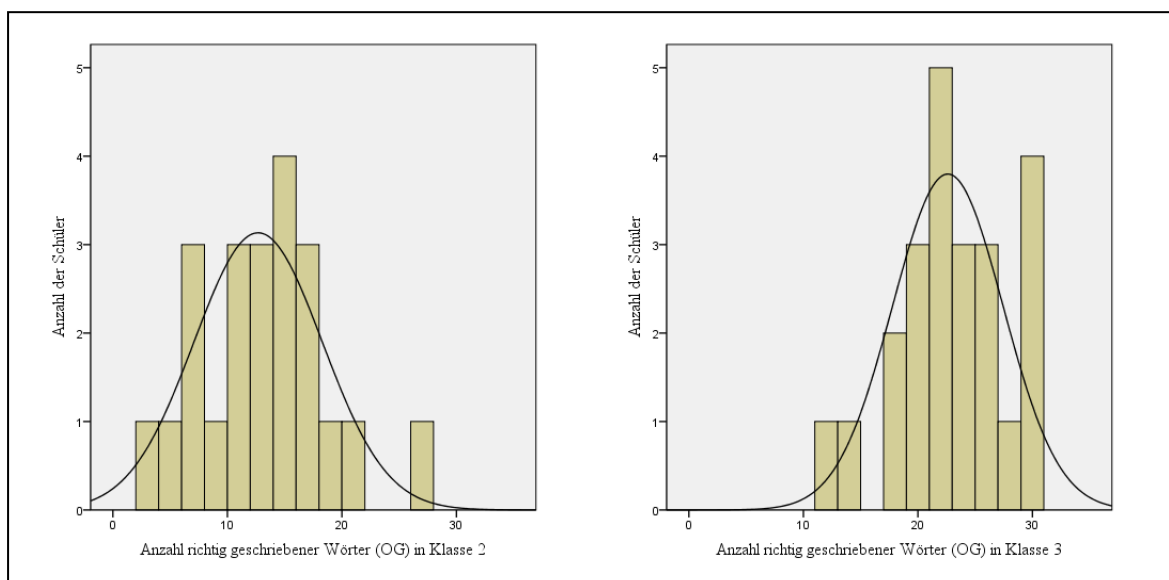


Abbildung 16: Häufigkeitsverteilung für den orthographischen Gesamtwert der Form A in Klasse 2 (N = 22) und in Klasse 3 (N = 23) im Rahmen der Erprobungsuntersuchung

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, OG = orthographischer Gesamtwert

Wie in Abbildung 16 zu erkennen, ist die Verteilungsform für den orthographischen Gesamtwert der Form A in Klasse 2 annähernd symmetrisch und eingipflig. Entsprechend

der statistischen Beurteilung auf Normalverteilung mit dem Shapiro-Wilk-Test liegt mit $p = .85$ ein nicht signifikanter Wert vor. Es handelt sich damit um eine Normalverteilung. In Klasse 2 wurden bei einer Standardabweichung von $SD = 6$ im Mittel 13 von 31 Wörtern richtig geschrieben. In Klasse 3 ist die Verteilungsform für den orthographischen Gesamtwert der Form A ebenfalls annähernd symmetrisch und eingipflig. Entsprechend der statistischen Beurteilung auf Normalverteilung mit dem Shapiro-Wilk-Test liegt mit $p = .41$ ein nicht signifikanter Wert vor. Es handelt sich damit um eine Normalverteilung. In Klasse 3 wurden bei einer Standardabweichung von $SD = 5$ im Mittel 23 von 31 Wörtern richtig geschrieben. Da der Mittelwert von der 2. zur 3. Klassenstufe steigt, ist davon auszugehen, dass die Schüler im Laufe eines Schuljahres im Rechtschreiben Lernfortschritte erzielen und die Form A diese erfassen kann. Tabelle 9 gibt die Mittelwerte und Standardabweichungen sowie das Minimum und Maximum für den orthographischen Gesamtwert der Form A in Klasse 2 und 3 im Rahmen der Erprobungsuntersuchung zusammenfassend wieder.

Tabelle 9: Kennwerte des orthographischen Gesamtwerts der Form A in Klasse 2 und 3 im Rahmen der Erprobungsuntersuchung

	N	M	SD	Min.	Max.
Klasse 2	22	12.7	5.6	3	26
Klasse 3	23	22.6	4.8	12	30

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, M = Mittelwert, SD = Standardabweichung, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilung für den orthographischen Gesamtwert der Form B ist in Abbildung 17 dargestellt.

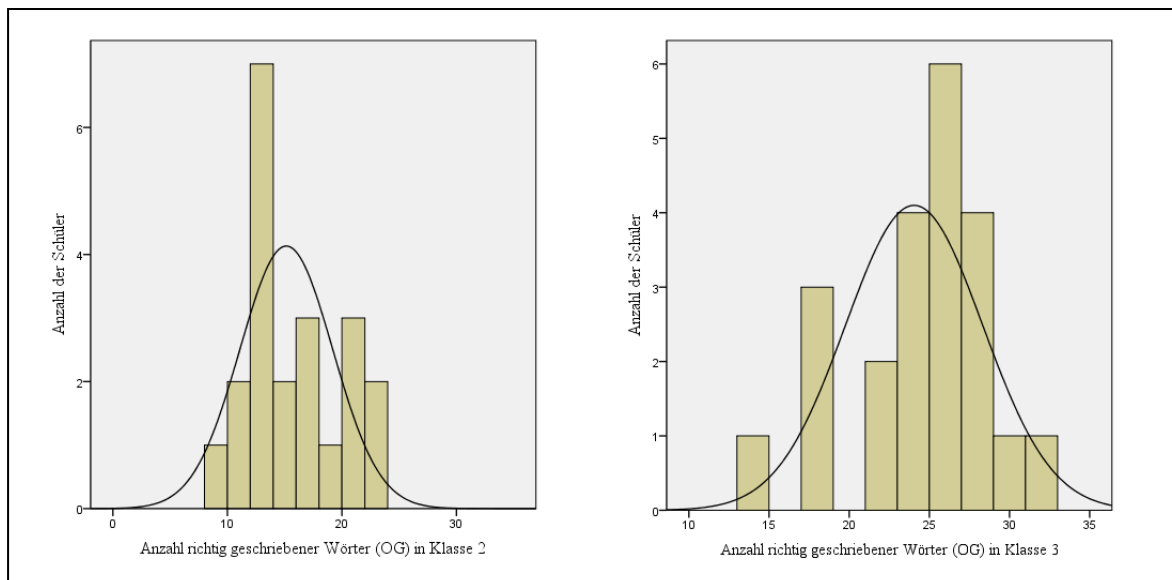


Abbildung 17: Häufigkeitsverteilung für den orthographischen Gesamtwert der Form B in Klasse 2 (N = 21) und in Klasse 3 (N = 22) im Rahmen der Erprobungsuntersuchung

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, OG = orthographischer Gesamtwert

Wie in Abbildung 17 zu erkennen, ist die Verteilungsform für den orthographischen Gesamtwert der Form B in Klasse 2 leicht asymmetrisch und eingipflig. Entsprechend der statistischen Beurteilung auf Normalverteilung mit dem Shapiro-Wilk-Test liegt mit $p = .20$ ein nicht signifikanter Wert vor. Es handelt sich damit um eine Normalverteilung. In Klasse 2 wurden bei einer Standardabweichung von $SD = 4$ im Mittel 15 von 31 Wörtern richtig geschrieben. In Klasse 3 ist die Verteilungsform für den orthographischen Gesamtwert der Form B annähernd symmetrisch und eingipflig. Entsprechend der statistischen Beurteilung auf Normalverteilung mit dem Shapiro-Wilk-Test liegt mit $p = .16$ ein nicht signifikanter Wert vor. Es handelt sich damit um eine Normalverteilung. In Klasse 3 wurden bei einer Standardabweichung von $SD = 4$ im Mittel 24 von 31 Wörtern richtig geschrieben. Da der Mittelwert von der 2. zur 3. Klassenstufe steigt, ist auch für die Form B davon auszugehen, dass die Schüler im Laufe eines Schuljahres im Rechtschreiben Lernfortschritte erzielen und die Form B diese erfassen kann. Tabelle 10 gibt die Mittelwerte und Standardabweichungen sowie das Minimum und Maximum für den orthographischen Gesamtwert der Form B in Klasse 2 und 3 im Rahmen der Erprobungsuntersuchung zusammenfassend wieder.

Tabelle 10: Kennwerte des orthographischen Gesamtwerts der Form B in Klasse 2 und 3 im Rahmen der Erprobungsuntersuchung

	N	M	SD	Min.	Max.
Klasse 2	21	15.1	4.0	9	23
Klasse 3	22	24.0	4.3	14	31

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, M = Mittelwert, SD = Standardabweichung, Min. = Minimum, Max. = Maximum

Der Vergleich zwischen den Mittelwerten der Form A und der Form B gibt Hinweise darauf, dass die Form B etwas leichter zu sein scheint als die Form A. Inhaltlich gibt es jedoch keine Hinweise darauf, dass die Form B leichter als die Form A ist, da die Wörter der Form A und B nach geschätzten gleichen Schwierigkeitsgraden ausgewählt wurden. Deshalb kann davon ausgegangen werden, dass die Mittelwertunterschiede der Formen A und B durch die kleine Anzahl der Erprobungsstichprobe bedingt sind.

Aufgrund der Ergebnisse der Erprobung und der Gespräche mit vier Grundschullehrkräften zeigte sich kein substanzieller Veränderungsbedarf der FE-RS 2 an. Im Lehrerheft wurden lediglich zwei Änderungen vorgenommen. Zum einen wurde die Angabe der Durchführungsdauer revidiert: Der Satz „Es ist davon auszugehen, dass die Durchführung einschließlich Vorbereitung ca. 20 Minuten dauert“ wurde verändert und lautet nun „Es ist davon auszugehen, dass die Durchführung einschließlich Vorbereitung am Anfang der 2. Klasse ca. 30 Minuten und am Ende der 2. Klasse ca. 15 Minuten dauert“. Zum anderen wurde eine wörtliche Instruktion für die Durchführung, nach dem Wort Nr. 28, hinzugefügt: „Bitte blättert jetzt um. Wir sind auf Seite 3“.

4.2.3 Entwicklung des Lehrerfragebogens zur Einschätzung der diagnostischen Verfahren zur Lernverlaufsdokumentation und Entwicklung des Lehrerfragebogens zur Einschätzung der Rechtschreibleistung der Schüler

Für die vorliegende Untersuchung wurde ein Lehrerfragebogen zur Einschätzung der diagnostischen Verfahren zur Lernverlaufsdokumentation, der FE-RS 2 und FE-AF 2 (Kuhlmann & Hartke, 2009a), entwickelt. Hinsichtlich der FE-RS 2 war das damit verbundene Ziel, Aussagen über die Praktikabilität, den pädagogischen Nutzen und das Verhalten der Schüler während der Testdurchführung zu erhalten. Zudem interessierten die im Unterricht verwendeten Methoden und Sozialformen sowie Verbesserungsvorschläge

und allgemeine Anmerkungen zur FE-RS 2. Der Lehrerfragebogen befindet sich im Anhang C.

Des Weiteren wurde für die vorliegende Untersuchung ein Lehrerfragebogen zur Einschätzung der Rechtschreibleistung der Schüler entwickelt. In diesem Fragebogen wurden die Lehrkräfte gebeten, die Rechtschreibleistungen ihrer Schüler in Form von Noten einzuschätzen. Die Entwicklung des Fragebogens orientierte sich an dem Bewertungssystem nach Ziffernnoten, da solche Noten, trotz ihrer Anfälligkeit für Fehler und Verzerrungen (Schabmann & B. M. Schmidt, 2009), ein gängiges Vorgehen zur Beurteilung schulischer Leistungen sind. Die Ergebnisse zur Einschätzung der schulischen Leistungen sollen anschließend für die Prüfung der externen Validität der FE-RS 2 herangezogen werden. Der Lehrerfragebogen zur Einschätzung der Rechtschreibleistung der Schüler befindet sich im Anhang D.

4.3 Stichprobengewinnung

Die Grundgesamtheit der vorliegenden Untersuchung bilden Grundschüler der 2. Klassenstufe. Da die Grundgesamtheit aus natürlichen Gruppen, den 2. Klassen, besteht, bietet es sich an, die Stichprobe für diese Untersuchung über eine Klumpenstichprobe zu ziehen. Eine Klumpenstichprobe besteht aus mehreren und zufällig ausgewählten, natürlich vorgruppierten Teilmengen, sogenannten Klumpen, der Grundgesamtheit (Bortz & Döring, 2006; Bortz & Schuster, 2010). Die Klumpenstichprobe wurde für die vorliegende Untersuchung auf das Bundesland Mecklenburg-Vorpommern und dessen Schulamtsbezirk Neubrandenburg mit dem Landkreis Mecklenburg-Strelitz und der Stadt Neubrandenburg eingegrenzt. Gründe dafür waren die begrenzten personellen und finanziellen Mittel der Untersuchung sowie die Bedingung, die Schulen innerhalb von maximal einer Stunde Fahrzeit zu erreichen. Innerhalb der ländlichen Region Mecklenburg-Strelitz und der städtischen Region Neubrandenburg wurden die Grundschulen und Regionalen Schulen mit Grundschulteil randomisiert ausgewählt.

Die Rekrutierung der Stichprobe erfolgte im August 2009. Es wurden alle 15 Grundschulen und 3 Regionalen Schulen mit Grundschulteil in Mecklenburg-Strelitz und Neubrandenburg per E-Mail mit der Bitte angeschrieben, die Untersuchung mit ihrer

Teilnahme zu unterstützen. Der Vordruck der E-Mail befindet sich im Anhang E. Die eigentliche Grundlage für die Rekrutierung der Stichprobe bildete § 71 des Schulgesetzes für das Land Mecklenburg-Vorpommern (MBWK, 2009b). Demnach müssen wissenschaftliche Forschungsvorhaben an Schulen öffentlicher Trägerschaft durch die zuständige Schulaufsichtsbehörde, hier das Schulamt Neubrandenburg, genehmigt werden. Im September 2009 wurde der Antrag auf Genehmigung der wissenschaftlichen Untersuchung an Grundschulen im Schulamt Neubrandenburg eingereicht. Der Antrag befindet sich im Anhang F. Der Antragsstellung im Schulamt folgte die Mitteilung über die Antragsstellung an alle Grundschulen und Regionalen Schulen mit Grundschulteil in Mecklenburg-Strelitz und Neubrandenburg. Die Mitteilung befindet sich im Anhang G. Alle Schulen, die auf die E-Mail und Mitteilung nicht reagierten, wurden telefonisch kontaktiert und gebeten, die Untersuchung mit ihrer Teilnahme zu unterstützen. Der Bitte mancher Schulen, die Untersuchung in einem persönlichen Gespräch vorzustellen, kam die Verfasserin der vorliegenden Arbeit im September und Oktober 2009 nach. Nachdem im Oktober 2009 der Untersuchung durch das Schulamt Neubrandenburg stattgegeben wurde, siehe Anhang H, wurden den 13 teilnehmenden Schulen Elterneinverständniserklärungen, siehe Anhang I, und die notwendigen Testunterlagen postalisch zugesandt.

4.4 Untersuchungsdurchführung und Datenauswertung

Die Durchführung der Untersuchung wurde durch die Verfasserin der vorliegenden Arbeit koordiniert. Die diagnostischen Verfahren wurden zu allen Messzeitpunkten der Untersuchung sowohl von externen Erhebern als auch von Lehrkräften durchgeführt. Externe Erheber waren neben der Verfasserin der vorliegenden Arbeit Studierende des Lehramts für Sonderpädagogik, des Lehramts an Grund- und Hauptschulen und an Gymnasien der Universität Rostock sowie Studierende des Studiengangs Early Education der Hochschule Neubrandenburg. Alle externen Erheber verfügten, bis auf Studierende des Lehramts für Sonderpädagogik, über sehr geringe diagnostische Kenntnisse. Deshalb wurden den Erhebern in gemeinsamen Treffen die wichtigsten theoretischen Grundlagen der Verfahren zur Lernverlaufsdokumentation und die Grundlagen zur Durchführung testdiagnostischer Verfahren erläutert. Zum 1. Messzeitpunkt wurden lediglich zwei Schulen externe Erheber zur Verfügung gestellt, da dies deren Voraussetzung für ihre Teilnahme war. In allen anderen Schulen führten Lehrkräfte die Verfahren durch. Ab dem

2. Messzeitpunkt wurden in der Hälfte beider Gruppen von Schulklassen, also in insgesamt 14 Klassen, externe Erheber zur Verfügung gestellt, um das Phänomen „teaching to the test“ (Helmke, 2007, S. 63) kontrollieren zu können. Die Gefahr dessen bestand, da die Lehrkräfte Inhalte und Anforderungsprofile der Verfahren zur Lernverlaufsdokumentation aus dem 1. Messzeitpunkt kannten und daraufhin bewusst oder unbewusst testvorbereitende Maßnahmen hätten ergreifen können. Externe Erheber wurden vor allem den Lehrkräften zur Seite gestellt, die nach dem 1. Messzeitpunkt den Wunsch nach Unterstützung äußerten.

Für die Durchführung der diagnostischen Verfahren wurden alle teilnehmenden Schulen eine Woche vor dem jeweiligen Messzeitpunkt und zu Beginn des jeweiligen Untersuchungszeitraumes per E-Mail mit dem Hinweis angeschrieben, die Verfahren entsprechend des Lehrerhefts mit den Durchführungshinweisen anzuwenden. Für die Durchführung der diagnostischen Verfahren wurden mit den Lehrkräften, denen externe Erheber zugeteilt waren, Termine vereinbart. Die Verfahren wurden im laufenden Unterricht, am frühen Vormittag, durchgeführt. Für jedes diagnostische Verfahren wurde von den Schulen eine Unterrichtsstunde bereitgestellt. Aufgrund der geringen Anzahl zur Verfügung stehender externer Erheber wurden alle diagnostischen Verfahren zumeist an einem Tag durchgeführt. Dabei wurden die Verfahren nach ihrem Anforderungsgrad, von schwer nach leicht, geordnet, um die hohe Belastung der Schüler und eine negative Beeinflussung auf die Testergebnisse möglichst gering zu halten. Dennoch war vor allem bei leistungsschwachen Schülern eine sinkende Aufmerksamkeit im Laufe des Tages zu beobachten. In vereinzelten Fällen wurden die Verfahren an zwei oder drei aufeinander folgenden Tagen durchgeführt. Lehrkräfte, die die Verfahren selbstständig in ihren Klassen durchführten, organisierten sich die Phase der Testdurchführung auf der Grundlage der vorgegebenen Durchführungshinweise eigenständig. Nach Abschluss jedes Messzeitpunktes wurden die Testunterlagen in den Schulen durch die Verfasserin der vorliegenden Arbeit abgeholt, um in ständigem Kontakt mit den Lehrkräften zu stehen und um die Untersuchung betreffende Fragen und Probleme zeitnah und persönlich zu klären.

Eine besondere Schwierigkeit in der Untersuchungsdurchführung, die den zeitlichen Verlauf der Untersuchung beeinflusste, war die Anwerbung von externen Erhebern. Vor dem Hintergrund, dass das Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation der Universität Rostock für die vorliegende Untersuchung keine personelle

Unterstützung bereitstellen konnte, war es anfänglich das Ziel, eine Kooperation zwischen der vorliegenden Untersuchung und dem Studiengang Early Education der Hochschule Neubrandenburg zu initiieren. Ziel sollte sein, die Untersuchungsdurchführung mit dortigen Praxisaufgaben der Studierenden zu verbinden. Eine Kooperation kam aufgrund der fortgeschrittenen Semesterplanungen aber nicht zustande. Letztlich konnten externe Erheber nur über den Freundeskreis der Verfasserin der vorliegenden Arbeit und das soziale Netzwerk StudiVZ angeworben werden, sodass für jeden Messzeitpunkt zwei bis fünf externe Erheber zur Verfügung standen. Aufgrund dieser geringen Anzahl externer Erheber verlängerte sich der Zeitraum des 1., 2. und 3. Messzeitpunktes auf knapp zwei Wochen. Weitere Gründe für die zeitliche Verlängerung von Messzeitpunkten waren Schulausfall aufgrund von Witterungsbedingungen, Krankheit von Lehrkräften und das schlichte Vergessen der Untersuchungsdurchführung durch Lehrkräfte.

Alle über die 4 Messzeitpunkte erhobenen Daten der 644 teilnehmenden Schüler wurden durch die Verfasserin dieser Arbeit in die Statistiksoftware Statistical Package for the Social Sciences (SPSS) eingegeben und ausgewertet. Für die Datenauswertung wurden jedoch nur Schüler mit vollständigen Datensätzen zu allen 4 Messzeitpunkten berücksichtigt, um die Problematik fehlender Daten bei statistischen Methoden der Datenanalyse zu umgehen. Die statistische Prüfung der FE-RS 2 erfolgte getrennt nach Form A und Form B. Aufgrund des Untersuchungsdesigns erfolgte die Prüfung auf prognostische Validität und Veränderungssensitivität der FE-RS 2 vom 1. zum 3. Messzeitpunkt sowie vom 2. zum 4. Messzeitpunkt, weil es sich dabei jeweils um die gleiche Teilstichprobe handelt. Auf eine Ergebnisdarstellung der ERO-Subskala wird in der vorliegenden Arbeit verzichtet, da diese Subskala nur eine Lupenstelle enthält und deshalb nicht zur differenzierten Analyse der Rechtschreibleistung beiträgt.

4.5 Forschungsfragen und Hypothesen

Nachfolgend werden die Forschungsfragen und Hypothesen (H) über die FE-RS 2 formuliert, die den Ausgangspunkt für die empirische Evaluation bilden.

Forschungsfrage 1:

Differenziert die FE-RS 2 zwischen Schülern mit guten und schlechten

Rechtschreibleistungen zu den 4 Messzeitpunkten?

Die erste Forschungsfrage zielt auf die Differenzierungsfähigkeit des Testverfahrens ab. Nach Lienert und Raatz (1998) ist eine gute Differenzierungsfähigkeit eines Tests abhängig von einer großen Standardabweichung der Messwerte bei hoher Reliabilität, um den Standardmessfehler niedrig zu halten. Des Weiteren nennen Lienert und Raatz (1998) die Symmetrie und Eingipfligkeit der Rohwertpunkteverteilung als Bedingungen für eine gute Differenzierungsfähigkeit. Standardabweichung, Reliabilitätsanalyse und Analyse der Häufigkeitsverteilung sollen neben der Itemschwierigkeit und Trennschärfe auch in der vorliegenden Arbeit als Maße für die Differenzierungsfähigkeit der FE-RS 2 herangezogen werden. Die Hypothesen zur ersten Forschungsfrage werden nachfolgend genannt.

H_{1,1}: Die Häufigkeitsverteilungen für den orthographischen Gesamtwert und die Subskalen der Form A und der Form B sind für alle 4 Messzeitpunkte glockenförmig, symmetrisch und eingipflig.

H_{1,2}: Die Standardabweichungen der Messwerte für den orthographischen Gesamtwert und die Subskalen der Form A und der Form B sind für alle 4 Messzeitpunkte hoch.

H_{1,3}: Für die Itemschwierigkeiten des orthographischen Gesamtwerts und der Subskalen der Form A und der Form B gilt zu allen 4 Messzeitpunkten $.2 \leq p_i \leq .8$.

H_{1,4}: Für die Trennschärfen des orthographischen Gesamtwerts und der Subskalen der Form A und der Form B gilt zu allen 4 Messzeitpunkten $.3 \leq r_{it} \leq .5$.

Forschungsfrage 2:

Ist die FE-RS 2 ein zuverlässiges Verfahren, das genau und exakt misst?

Die zweite Forschungsfrage zielt auf die Reliabilität des Testverfahrens ab. Zur Reliabilitätsschätzung der FE-RS 2 sollen die Daten zur Retest-, und Testhalbierungsreliabilität sowie die interne Konsistenz herangezogen werden. Die Hypothesen zur zweiten Forschungsfrage werden nachfolgend genannt.

H_{2,1}: Für die Retestreliabilität des orthographischen Gesamtwerts der Form A und Form

B gilt $r_{tt} > .70$. Für die Retestreliabilität der Subskalen der Form A und Form B gilt $.50 < r_{tt} \leq .70$.

H_{2,2}: Für die Testhalbierungsreliabilität des orthographischen Gesamtwerts der Form A und der Form B gilt $r_{tt} > .70$. Für die Testhalbierungsreliabilität der Subskalen der Form A und Form B gilt $.50 < r_{tt} \leq .70$.

H_{2,3}: Für die interne Konsistenz des orthographischen Gesamtwerts der Form A und der Form B gilt $r_{tt} > .70$. Für die interne Konsistenz der Subskalen der Form A und Form B gilt $.50 < r_{tt} \leq .70$.

Forschungsfrage 3:

Misst die FE-RS 2 die Rechtschreibleistung der Zweitklässler?

Die dritte Forschungsfrage zielt auf die Validität des Testverfahrens ab. Zur Bestimmung der Validität der FE-RS 2 sollen die Inhaltsvalidität, die Übereinstimmungsvalidität mit der Rechtschreibnote, die konkurrente Validität mit den Testergebnissen des DERET 1-2+ (Stock & Schneider, 2008a) und der FE-L 2 (Kuhlmann & Hartke, 2009b) sowie die diskriminante Validität mit den Testergebnissen des DEMAT 1+ (Krajewski et al., 2002), des DEMAT 2+ (Krajewski et al., 2004) und der FE-AF 2 (Kuhlmann & Hartke, 2009a) sowie die faktorielle Validität herangezogen werden. Die Hypothesen zur dritten Forschungsfrage werden nachfolgend genannt.

H_{3,1}: Die FE-RS 2 ist inhaltsvalide.

H_{3,2}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und dem quantitativen Testergebnis des DERET 1-2+ (Stock & Schneider, 2008a) zum 1. und 4. Messzeitpunkt gilt $r > .70$.

H_{3,3}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und der Anzahl richtig gelesener Wörter der FE-L 2 (Kuhlmann & Hartke, 2009b) zum 1. und 4. Messzeitpunkt gilt $.50 < r \leq .70$.

H_{3,4}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A

sowie der Form B und dem quantitativen Testergebnis der FE-AF 2 (Kuhlmann & Hartke, 2009a) zum 1. und 4. Messzeitpunkt gilt $r \leq .50$.

H_{3,5}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und dem Testergebnis des DEMAT 1+ (Krajewski et al., 2002) zum 1. Messzeitpunkt und des DEMAT 2+ (Krajewski et al., 2004) zum 4. Messzeitpunkt gilt $r \leq .50$.

H_{3,6}: Die Form A und Form B erfassen die Subskalen ELG, ERG, ELO und ERO. Es zeigen sich stabile Zuordnungen der Items der Subskalen zu Faktoren über alle 4 Messzeitpunkte.

H_{3,7}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und den Rechtschreibnoten der Schüler zum 4. Messzeitpunkt gilt $r > .70$.

Forschungsfrage 4:

Kann aufgrund der Daten des 1. und 2. Messzeitpunktes die Rechtschreibleistung der Zweitklässler für den 3. und 4. Messzeitpunkt vorausgesagt werden?

Die vierte Forschungsfrage zielt auf die prognostische Aussagekraft der FE-RS 2 ab. Zur Schätzung der prognostischen Validität der FE-RS 2 sollen Regressionsanalysen und klassifikatorische Analysen herangezogen werden. Die Hypothesen zur vierten Forschungsfrage werden nachfolgend genannt.

H_{4,1}: Der orthographische Gesamtwert und die Subskalen der Form A und der Form B zum 1., 2. und 3. Messzeitpunkt tragen signifikant zur Varianzaufklärung des orthographischen Gesamtwerts zum 2., 3. und 4. Messzeitpunkt bei. Für den RATZ-Index gilt $> 66 \%$.

H_{4,2}: Der orthographische Gesamtwert und die Subskalen der Form A und der Form B zum 1., 2. und 3. Messzeitpunkt sagen die Rechtschreibnote zum 4. Messzeitpunkt vorher. Für den RATZ-Index gilt $> 66 \%$.

Forschungsfrage 5:

Handelt es sich bei den Formen A und B der FE-RS 2 um Parallelversionen des Verfahrens?

Die fünfte Forschungsfrage zielt auf die Äquivalenz beider Formen ab. Nach Lienert und Raatz (1998) umfasst die Äquivalenz von Parallelformen eines Verfahrens hinreichend gleiche Varianzen und Mittelwerte, äquivalente Häufigkeitsverteilungen, hinreichend gleiche Reliabilitäten und eine hohe Paralleltestreliabilität sowie hinreichend gleiche Validitäten. Die genannten Äquivalenzkriterien sollen auch in der vorliegenden Arbeit als Kennwerte für Äquivalenz herangezogen werden. Die Hypothesen zur fünften Forschungsfrage werden nachfolgend genannt.

H_{5.1}: Die Varianzen und Mittelwerte zwischen den orthographischen Gesamtwerten, den Subskalen der Form A und Form B zu den 4 Messzeitpunkten unterscheiden sich nicht signifikant.

H_{5.2}: Die Häufigkeitsverteilungen der Form A und Form B unterscheiden sich nicht voneinander.

H_{5.3}: Für die Paralleltestreliabilität des orthographischen Gesamtwerts und der Subskalen der Form A und Form B gilt $r_{tt} > .70$.

H_{5.4}: Die Reliabilitätskoeffizienten der Form A und Form B unterscheiden sich nicht signifikant.

H_{5.5}: Die Validitätskoeffizienten der Form A und Form B unterscheiden sich nicht signifikant.

Forschungsfrage 6:

Erfasst die FE-RS 2 Veränderungen der Rechtschreibleistungen in der 2. Klassenstufe zu den 4 Messzeitpunkten?

Die sechste Forschungsfrage zielt auf die Veränderungssensitivität des Testverfahrens ab. Nach Bortz und Döring (2006) umfasst die Veränderungssensitivität eines Tests hinreichend einfache Differenzmaße, die auch in der vorliegenden Arbeit als Maße für die

Veränderungssensitivität der FE-RS 2 herangezogen werden. Die Hypothese zur sechsten Forschungsfrage wird nachfolgend genannt.

H_{6.1}: Die Mittelwerte des orthographischen Gesamtwerts und der Subskalen der Form A und der Form B steigen signifikant vom 1. zum 3. Messzeitpunkt und vom 2. zum 4. Messzeitpunkt.

4.6 Statistische Methoden der Datenanalyse

Eine Veränderungsmessung erfasst Unterschiede und Änderungen von Zuständen und Bedingungen (Gollwitzer & Jäger, 2009; Petermann, 2000). „Die Analyse von Veränderungen zählt zu den interessantesten, aber auch schwierigsten Aufgaben der Human- und Sozialwissenschaften“ (Bortz & Döring, 2006, S. 547), denn zum einen führt die Veränderungsmessung mit Testverfahren, die auf der KTT basieren, zu Schwierigkeiten und zum anderen ist die Veränderungsmessung mit Testverfahren, die auf der PTT basieren, mit einem sehr großen Konstruktions- und Evaluationsaufwand verbunden. Nachfolgend werden die testtheoretischen Grundlagen der KTT und PTT überblicksartig erläutert und im Hinblick auf Möglichkeiten der Veränderungsmessung bewertet.

Die KTT ist eine Messfehlertheorie, die annimmt, dass sich das Testergebnis des untersuchten Merkmals aus einem *wahren Ausprägungsgrad* des untersuchten Merkmals und einem *zufälligen Messfehler* zusammensetzt. Das Testergebnis entspricht damit, abgesehen vom Messfehler, der Merkmalsausprägung. Die KTT schließt auf die wahre Merkmalsausprägung, indem der Messfehler ermittelt wird. Um auf die wahre Merkmalsausprägung schließen zu können, wird sich auf folgende *fünf Axiome der KTT* berufen:

1. Das Testergebnis setzt sich aus einem wahren Wert und einem Messfehler zusammen.
2. Bei Messwiederholungen bei selben Personen ändert sich nicht der wahre Ausprägungsgrad des untersuchten Merkmals.
3. Die Höhe des Messfehlers ist unabhängig vom wahren Ausprägungsgrad des untersuchten Merkmals.
4. Die Höhe des Messfehlers ist unabhängig vom Ausprägungsgrad anderer untersuchter

Merkmale.

5. Die Messfehler bei Messwiederholungen oder bei verschiedenen Personen sind unabhängig voneinander (Bortz & Döring, 2006).

Entsprechend der Axiome geht die KTT bei Messwiederholungen von einer zeitlichen Stabilität der Merkmale aus. Veränderungen werden in der KTT demzufolge als Messfehler interpretiert. In Hinblick auf Möglichkeiten der Veränderungsmessung zeigen sich im Rahmen der KTT damit einige Widersprüche und Schwierigkeiten. Zum Ersten basiert die Anwendung von Verfahren zur Lernverlaufsmessung auf der Annahme, dass sich Merkmale über einen gewissen Zeitraum verändern und nicht stabil sind. Ansonsten wären mehrmalige Messwiederholungen und Interventionen nicht sinnvoll. Zum Zweiten können sich neben wahren Veränderungswerten auch äußere Störeinflüsse, Testeffekte oder Regressionseffekte im Testergebnis niederschlagen, wobei unklar ist, ob es sich um Messfehler oder um wahre Veränderungen handelt (Clasen, 2010). Zum Dritten ist das Reliabilitäts-Validitäts-Dilemma unvereinbar mit der Veränderungsmessung im Rahmen der KTT. Das Dilemma entsteht dadurch, dass die Reliabilität der Differenzwerte umso niedriger ist, je höher die Korrelation zwischen zwei Messzeitpunkten ist (Rentzsch & Schütz, 2009). „Obwohl man beides in der Prozessdiagnostik erhalten möchte – hohe Reliabilitäten der Differenzwerte und hohe Validitäten der Einzelmessungen – scheint das aufgrund der Implikationen der KTT nicht realisierbar“ (Rentzsch & Schütz, 2009, S. 228).

Eine Alternative, Veränderungsmessungen zu erheben und zu analysieren, bietet die PTT. Die PTT nimmt an, dass das Testergebnis ein *Indikator für latente Dimensionen* des untersuchten Merkmals ist. Das Testergebnis entspricht damit nicht wie in der KTT der direkten Merkmalsausprägung, sondern der Wahrscheinlichkeit, mit der eine bestimmte Aufgabenantwort von einer latenten Merkmalsdimension abhängt (Bortz & Döring, 2006). Ein Beispiel; ein Zweitklässler mit besserer Rechtschreibfähigkeit schreibt das Wort *Fahrräder* mit höherer Wahrscheinlichkeit orthographisch richtig als ein Zweitklässler mit schlechterer Rechtschreibfähigkeit. Arnold und Hartig (2008) resümieren in Anlehnung an Rauch und Hartig (2007), dass die PTT im Gegensatz zur KTT Vorteile bei der Veränderungsmessung hat, denn PTT-basierte Auswertungsmethoden erlauben, Veränderungen von Merkmalen zu untersuchen. Somit können z. B. Aussagen darüber getroffen werden, welche Aufgaben eine Person mit welcher Sicherheit lösen kann oder welche Aufgaben zu unterschiedlichen Messzeitpunkten für eine Person zu schwer waren bzw. gelöst werden können. Ein weiterer Vorteil der PTT ist, dass Fähigkeitsschätzungen

auf unterschiedlichen Aufgaben einer Aufgabengrundgesamtheit beruhen können und damit bei mehrmaligen Messwiederholungen Erinnerungseffekte vermieden werden (Arnold & Hartig, 2008).

Aus der Darstellung der testtheoretischen Grundlagen der KTT und PTT sowie deren Möglichkeiten für Veränderungsmessungen ergibt sich für die vorliegende Arbeit Folgendes: Da die KTT die Forschungslandschaft und damit die Forschungsliteratur sowie die zur Verfügung stehenden Statistikprogramme dominiert, wurden die Daten der vorliegenden Arbeit aus pragmatischen Gründen nach der KTT analysiert. Die statistische Datenanalyse erfolgte mit SPSS der Version 19. Aufgrund dessen werden im Folgenden die Grundbegriffe und die verwendeten statistischen Methoden der KTT der Datenanalyse erläutert.

Mittelwert

Der *Mittelwert*, auch arithmetisches Mittel oder Durchschnitt genannt, beschreibt das Zentrum einer Verteilung (Bortz & Schuster, 2010).

Varianz

Die *Varianz* ist ein Maß für die Streuung der Messwerte um den Mittelwert. Da die Varianz schwierig zu interpretieren ist, wird sie vorwiegend zur Berechnung der Standardabweichung und des Standardmessfehlers verwendet. Der Vergleich zweier Stichprobenvarianzen kann mit dem *Levene-Test* erfolgen (Bortz & Schuster, 2010).

Standardabweichung

Die *Standardabweichung* (SD), auch Streuung genannt, ist ebenfalls ein Maß für die Streuung der Messwerte um den Mittelwert (Bortz & Schuster, 2010). Sie beschreibt „die relative Position einzelner Messwerte im Vergleich zur Gesamtverteilung der Messwerte um den Mittelwert“ (Bundschuh, 2005, S. 95).

Standardmessfehler

Der *Standardmessfehler* (SE) ist ein Unsicherheitsmaß. Mit der Berechnung des Standardmessfehlers kann der Bereich bestimmt werden, in dem die wahre Merkmalsausprägung eines untersuchten Merkmals liegt. Dieser Bereich wird auch *Vertrauensbereich*, Vertrauensintervall oder Konfidenzintervall genannt. Der

Standardmessfehler wird wie folgt berechnet: $SE = SD \cdot \sqrt{(1-r_{tt})}$; (r_{tt} = Reliabilitätskoeffizient) (Bundschuh, 2005).

Signifikanztest und Irrtumswahrscheinlichkeit

Der *Signifikanztest* ist ein Verfahren zur statistischen Hypothesenprüfung, das überprüft, wie gut die Stichprobenergebnisse zur Nullhypothese passen. Das bedeutet, es wird überprüft, ob die Nullhypothese akzeptiert und die Alternativhypothese abgelehnt oder ob die Nullhypothese abgelehnt und die Alternativhypothese akzeptiert werden kann. Um diese Entscheidung treffen zu können, wird die *Irrtumswahrscheinlichkeit* p ermittelt (Bortz & Döring, 2006). Die Irrtumswahrscheinlichkeit gibt an, „mit welcher bedingten Wahrscheinlichkeit das gefundene Untersuchungsergebnis [Stichprobenergebnis] auftritt, wenn in der Population die Nullhypothese gilt“ (Bortz & Döring, 2006, S. 25). Eine hohe Irrtumswahrscheinlichkeit bedeutet, dass es keine statistisch bedeutsamen Unterschiede/Zusammenhänge/Veränderungen zwischen dem Stichprobenergebnis und der Population gibt. Bei einem nicht signifikanten Ergebnis wird die Nullhypothese demzufolge akzeptiert und die Alternativhypothese abgelehnt. Eine niedrige Irrtumswahrscheinlichkeit bedeutet, dass es statistisch bedeutsame Unterschiede/Zusammenhänge/Veränderungen zwischen dem Stichprobenergebnis und der Population gibt. Bei einem signifikanten Ergebnis wird die Nullhypothese demzufolge abgelehnt und die Alternativhypothese akzeptiert (Bortz & Döring, 2006). Die in der Forschungspraxis gängige Klassifikation von signifikanten und nicht signifikanten Ergebnissen wird in Tabelle 11 wiedergegeben.

Tabelle 11: Kriterien für signifikante und nicht signifikante Ergebnisse (Bühl, 2008, S. 121)

Irrtumswahrscheinlichkeit	Bedeutung	Symbolisierung
$p > 0.05$	nicht signifikant	ns
$p \leq 0.05$	signifikant	*
$p \leq 0.01$	sehr signifikant	**
$p \leq 0.001$	höchst signifikant	***

Anmerkungen: p = Irrtumswahrscheinlichkeit

Korrelation

Die *Korrelation* beschreibt den Zusammenhang zwischen Variablen. Das bedeutet, es wird überprüft, ob die Ausprägung einer Variable in Zusammenhang mit der Ausprägung einer

weiteren Variable steht. Um die Enge und Richtung eines Zusammenhangs beschreiben zu können, wird der *Korrelationskoeffizient* r berechnet. Sind die Variablen normalverteilt, wird die Korrelation nach Pearson berechnet. Sind die Variablen nicht normalverteilt, wird die Korrelation nach Spearman berechnet. Der Korrelationskoeffizient r hat einen Wertebereich von $-1 \leq r \leq 1$. Je größer der Korrelationskoeffizient ist, desto enger ist der Zusammenhang. Die Richtung eines Zusammenhangs kann positiv oder negativ sein. Eine positive Korrelation bedeutet, dass der Zusammenhang zwischen Variablen gleichsinnig ist, das heißt, je mehr, desto mehr bzw. je weniger, desto weniger. Eine negative Korrelation bedeutet, dass der Zusammenhang zwischen Variablen gegensinnig ist, das heißt, je mehr, desto weniger bzw. je weniger, desto mehr (Bortz & Döring, 2006; Bühner, 2004). Die übliche Einteilung der Größen der Korrelationskoeffizienten ist in Tabelle 12 dargestellt.

Tabelle 12: Einteilung der Größen der Korrelationskoeffizienten (Bühl, 2008, S. 346)

Wert	Interpretation
bis 0,2	sehr geringe Korrelation
bis 0,5	geringe Korrelation
bis 0,7	mittlere Korrelation
bis 0,9	hohe Korrelation
über 0,9	sehr hohe Korrelation

Die Prüfung der Signifikanz von Korrelationsunterschieden zweier unabhängiger Stichproben erfolgt über die *Prüfgröße* z , die nach folgender Gleichung berechnet wird: $z = Z_1 - Z_2 / \sqrt{((1 / n_1 - 3) + (1 / n_2 - 3))}$ (Anmerkungen: Z_1/Z_2 = transformierter Korrelationskoeffizient der Stichprobe 1/2 anhand der Tabelle F 9 (Bortz & Döring, 2006), n_1/n_2 = Stichprobenumfang der Stichprobe 1/2) (Bortz, 2005).

Häufigkeitsverteilung

Die *Häufigkeitsverteilung* gibt an, wie häufig eine Merkmalsausprägung für ein untersuchtes Merkmal in der Stichprobe vorkommt. Die Häufigkeitsverteilung beschreibt, unterstützt durch die grafische Darstellung z. B. in Form eines Histogramms oder Boxplots, die Verteilungsform der Stichprobe (Bortz & Schuster, 2010). Verfahren zur Überprüfung der Normalverteilung, die in der vorliegenden Arbeit angewandt werden, sind der *Kolmogorov-Smirnov-Test* (KST) und der *Shapiro-Wilk-Test*. Beide Tests prüfen die Nullhypothese, dass die Merkmale normalverteilt sind. „Für große Stichproben ist der

Kolmogorov-Smirnov-Test gut geeignet; liegen dagegen nur wenige Beobachtungen (etwa weniger als 50) vor, so sollte zumindest ergänzend der Shapiro-Wilks-Test herangezogen werden“ (Brosius, 2011, S. 405).

Mittelwertvergleich

Der *Mittelwertvergleich* ist ein Verfahren, das Mittelwerte von Stichproben miteinander vergleicht. Es wird überprüft, ob Mittelwertunterschiede signifikant sind. Für den Mittelwertvergleich zwischen zwei unabhängigen, normalverteilten Gruppen wird der *t-Test* verwendet, für den Mittelwertvergleich zwischen zwei unabhängigen, nicht normalverteilten Gruppen der *U-Test* von Mann-Whitney. Für den Mittelwertvergleich zwischen zwei abhängigen, nicht normalverteilten Gruppen wird der *Wilcoxon-Test* verwendet (Bortz & Döring, 2006; Bortz & Schuster, 2010; Bühl, 2008).

Itemschwierigkeit

Die *Itemschwierigkeit* ist ein Maß für die Lösungsrate von Items. Um die Itemschwierigkeit beschreiben zu können, wird der *Schwierigkeitsindex* p_i berechnet. Der Schwierigkeitsindex p_i hat einen Wertebereich von $0 \leq p_i \leq 1$. Je größer der Schwierigkeitsindex ist, desto leichter ist das Item und je mehr Probanden kommen zum richtigen Ergebnis (Bortz & Döring, 2006). Ein Schwierigkeitsindex von $p_i > .80$ gilt als niedrig, von $.80 \geq p_i \geq .20$ als mittel und von $p_i < .20$ als hoch (Fisseni, 2004).

Trennschärfe

Die *Trennschärfe* gibt an, „wie gut ein einzelnes Item das Gesamtergebnis eines Tests repräsentiert. Die Trennschärfe wird für jedes Item eines Tests berechnet und ist definiert als die Korrelation der Beantwortung dieses Items mit dem Gesamttestwert“ (Bortz & Döring, 2006, S. 219). Das bedeutet, dass Probanden mit einem hohen Gesamtergebnis auch trennscharfe Items richtig lösen und Probanden mit einem niedrigen Gesamtergebnis nicht. Ein trennscharfes Item kann also Probanden mit hoher und niedriger Merkmalsausprägung voneinander trennen. Der Trennschärfekoeffizient r_{it} hat einen Wertebereich von $-1 \leq r_{it} \leq 1$. Liegt die Trennschärfe nahe bei 1, misst das einzelne Item Ähnliches wie der Gesamttest und wird von Probanden mit hoher Merkmalsausprägung gelöst. Liegt die Trennschärfe nahe bei 0, gibt es keinen Zusammenhang zwischen dem einzelnen Item und dem Gesamttest und das Item kann nicht zwischen Probanden mit hoher und niedriger Merkmalsausprägung differenzieren. Liegt die Trennschärfe in der

Nähe von -1, wird das Item von Probanden mit niedriger Merkmalsausprägung gelöst (Bortz & Döring, 2006; Kelava & Moosbrugger, 2007). Eine Trennschärfe von $r_{it} < .30$ gilt als niedrig, von $.30 \leq r_{it} \leq .50$ als mittel und von $r_{it} > .50$ als hoch (Fisseni, 2004).

Reliabilität

Die *Reliabilität* gibt die Messgenauigkeit eines Messinstruments an. Im Rahmen der KTT ist die Reliabilität definiert „als Anteil der wahren Varianz [Varianz der wahren, messfehlerfreien Werte] . . . an der beobachteten Varianz [Varianz der tatsächlichen, messfehlerbehafteten Werte]“ (Bortz & Döring, 2006, S. 196). Um die Messgenauigkeit eines Verfahrens schätzen zu können, wird der *Reliabilitätskoeffizient* r_{tt} berechnet. Der Reliabilitätskoeffizient hat einen Wertebereich von $0 \leq r_{tt} \leq 1$. Je größer der Reliabilitätskoeffizient ist, desto kleiner ist die Fehlervarianz und desto genauer misst ein Testverfahren den wahren Wert (Bortz & Döring, 2006). Eine Reliabilität von $r_{tt} < .50$ gilt als niedrig, von $.50 \leq r_{tt} \leq .70$ als mittel und von $r_{tt} > .70$ als hoch (Bühl, 2008). Es gibt vier Methoden der Reliabilitätsschätzung, die nachfolgend kurz erläutert werden. Zur Schätzung der *Retestreliabilität*, auch Testwiederholungsmethode genannt, wird dasselbe Testverfahren derselben Stichprobe zweimal vorgelegt und die Korrelation zwischen beiden Messwertreihen berechnet. Zur Schätzung der *Paralleltestreliabilität* werden Parallelförmungen eines Testverfahrens derselben Stichprobe vorgelegt und die Korrelation zwischen beiden Formen berechnet. Zur Schätzung der *Testhalbierungsreliabilität*, auch Split-half-Reliabilität oder Äquivalenz genannt, werden die Items des Testverfahrens in zwei Testhälften mit annähernd gleicher Trennschärfe und gleichem Schwierigkeitsgrad geteilt. Es wird die Korrelation zwischen beiden Testhälften berechnet und mittels der Spearman-Brown-Formel von der Korrelation der beiden Testhälften auf die Gesamttestlänge korrigiert. Zur Schätzung der *internen Konsistenz* wird ein Testverfahren mit n Items in n Teile geteilt und die Korrelation zwischen den Items berechnet. Die Berechnung der internen Konsistenz erfolgt am häufigsten mittels des Alphakoeffizienten nach Cronbach (Bortz & Döring, 2006; Bundschuh, 2005; Schermelleh-Engel & Werner, 2007).

Validität

Die *Validität* gibt an, ob ein Test das misst, was er messen soll bzw. vorgibt zu messen. Es gibt drei Arten, um die Validität eines Testverfahrens zu bestimmen: *Inhaltsvalidität*, *Kriteriumsvalidität* und *Konstruktvalidität*. Bei der Bestimmung der Inhaltsvalidität wird

überprüft, ob das Testverfahren inhaltlich das zu messende Merkmal erfasst. Die Inhaltsvalidität wird nicht numerisch bestimmt, sondern erfolgt aufgrund fachlicher Einschätzungen. Bei der Bestimmung der Kriteriumsvalidität wird überprüft, ob das Testergebnis mit mindestens einem Kriterium korreliert, mit dem es aufgrund seines Messanspruchs in Zusammenhang stehen sollte. Das Kriterium kann zum selben Messzeitpunkt wie der Test (*Übereinstimmungsvalidität*), später (*prognostische Validität*) oder früher (*retrospektive Validität*) erhoben werden (Bühner, 2004). Die *inkrementelle Validität*, die auch der Kriteriumsvalidität zugeordnet ist, „bezeichnet den Beitrag eines Tests zur Verbesserung der Vorhersage eines Kriteriums über einen anderen Test hinaus“ (Bühner, 2004, S. 31). Bei der Bestimmung der Konstruktvalidität wird überprüft, „wie gut ein Test mit der Theorie übereinstimmt, von der seine Testkonstruktion ausging“ (Bundschuh, 2005, S. 89). Die Konstruktvalidität wird in *konkurrente*, *diskriminante* und *faktorielle Validität* unterschieden. Bei der konkurrenten Validität wird ein Test mit einem Test gleichen Messanspruchs, bei der diskriminanten Validität unterschiedlichen Messanspruchs korreliert. Die faktorielle Validität dient dazu, inhaltlich gleiche Items eines Tests zu einem Faktor zusammenzufassen und diese von anderen zu trennen (Bühner, 2004).

Dimensionalität und Faktorenanalyse

Die *Dimensionalität* gibt an, ob ein Testverfahren „nur *ein* Merkmal bzw. Konstrukt erfasst (eindimensio[sic]naler Test), oder ob mit den Testitems *mehrere* Konstrukte bzw. Teilkonstrukte operationalisiert werden (mehrdimensionaler Test)“ (Bortz & Döring, 2006, S. 221). Die Dimensionalitätsüberprüfung erfolgt mit der Faktorenanalyse. Die Faktorenanalyse ist ein datenreduzierendes Verfahren, das eine große Anzahl korrelierender Variablen zu einer kleineren Anzahl von Faktoren zusammenfasst (Bortz & Döring, 2006; Bühl, 2008). „Ein Faktor umfasst inhaltlich das Gemeinsame der zu ihm gehörenden korrelierenden Variablen“ (Bortz & Döring, 2006, S. 727). Es wird zwischen der *explorativen* und *konfirmativen Faktorenanalyse* unterschieden. Die explorative Faktorenanalyse (EFA) ist ein hypothesengenerierendes und struktursuchendes Verfahren. Ziel der EFA ist die Datenreduktion und die überschaubare Analyse und Interpretation der ermittelten Faktoren. Die konfirmative Faktorenanalyse (KFA) ist ein hypothesenprüfendes und strukturüberprüfendes Verfahren. Ziel der KFA ist, zu überprüfen, ob die erhobenen empirischen Daten mit dem zugrunde liegenden theoretischen Modell übereinstimmen (Bühl, 2008; Moosbrugger & Schermelleh-Engel,

2007). Als Voraussetzung für die Durchführung einer Faktorenanalyse sollte die Korrelationsmatrix auf Eignung mittels folgender Verfahren überprüft werden:

1. *Bartlett-Test* auf Sphärizität, der von der Nullhypothese ausgeht, dass alle Itemkorrelationen in der Population gleich Null sind,
2. *Kaiser-Meyer-Olkin-Koeffizient* (KMO-Koeffizient), der die gemeinsame Varianz aller Variablen bestimmt und $\geq .70$ sein sollte und
3. *Measure of Sample Adequacy-Koeffizient* (MSA-Koeffizient), der ähnlich wie der KMO-Koeffizient ist und ebenfalls $\geq .70$ sein sollte (Bühner, 2004).

Regression

Die *Regression* dient dazu, Merkmalsausprägungen vorherzusagen. Es ist zwischen der *linearen* und der *multiplen Regression* zu unterscheiden. Das Ziel der linearen Regression ist die Vorhersage einer abhängigen Variable y (Kriterium) durch eine unabhängige Variable x (Prädiktor). „Die lineare Regression geht von einem linearen Zusammenhang zwischen Prädiktor und Kriterium aus, so dass [sic] die vorausgesagten (geschätzten) Werte auf einer Geraden darstellbar sind“ (Leonhart, 2009, S. 284). Je weniger die geschätzten Werte von der Regressionsgeraden abweichen, desto sicherer ist die Vorhersage. Das heißt, eine Vorhersage ist dann sicher, wenn die Abweichung (Vorhersagefehler/Residuum) des tatsächlichen y -Wertes vom geschätzten y -Wert möglichst klein ist. Die Grundgleichung der linearen Regression lautet: $\hat{y}_i = b_{yx} \cdot x_i + a_{yx}$ (Anmerkungen: \hat{y}_i = Kriterium, b_{yx} = Regressionskoeffizient und Steigung der Geraden, x_i = Prädiktor und x -Wert der Person i , a_{yx} = additive Konstante und y -Achsen-Abschnitt) (Leonhart, 2009). Die Voraussetzungen für die Durchführung einer linearen Regression sind:

1. Normalverteilung und Intervallskalierung für Prädiktor und Kriterium,
2. Homoskedastizität (Varianz der Vorhersagefehler ist gleich [Bühner & Ziegler, 2009]),
3. Normalverteilung der Residuen,
4. Unabhängigkeit der Residuen (keine Autokorrelation) (Leonhart, 2009).

Das Ziel der multiplen Regression ist die Vorhersage eines Kriteriums durch mehrere Prädiktoren. Da die multiple Regression eine Erweiterung der linearen Regression ist, wird wie dort eine sichere Vorhersage mit möglichst kleinem Vorhersagefehler angestrebt. Die Gleichung der multiplen Regression lautet: $\hat{y}_i = b_1 \cdot x_{i1} + b_2 \cdot x_{i2} + \dots + b_k \cdot x_{ik} + a_{123\dots k}$ (Leonhart, 2009). Die Voraussetzungen für die Durchführung einer multiplen Regression sind:

1. lineare Zusammenhänge zwischen Prädiktoren sowie zwischen Prädiktoren und Kriterium,
2. Homoskedastizität,
3. Normalverteilung der Residuen,
4. Unabhängigkeit der Residuen (Autokorrelationen werden mittels des *Durbin-Watson-Koeffizienten* d überprüft, der einen Wertebereich von $0 \leq d \leq 4$ hat. Nach Bühl [2008] ist ein Wert in der Nähe von $d = 2$ akzeptabel.)
5. Aufnahme aller relevanten Prädiktoren, die mit dem Kriterium korrelieren, in die Regressionsanalyse,
6. keine perfekte oder sehr hohe Korrelation von zwei oder mehr Prädiktoren (Multikollinearität wird mittels des *Konditionsindex* CI überprüft. Ein Konditionsindex von > 15 deutet auf mäßige Kollinearität, ein Konditionsindex von > 30 auf starke Kollinearität hin.),
7. hohe Reliabilität der Prädiktoren und des Kriteriums,
8. keine eingeschränkten Varianzen der verwendeten Variablen (Bühner & Ziegler, 2009).

Klassifikatorische Analyse

Die *klassifikatorische Analyse* dient der individuellen Vorhersage von Merkmalsausprägungen und der Zuordnung von Schülern in eine Risikogruppe und Nicht-Risikogruppe. Die Schüler der Risikogruppe sollten entwicklungsgefährdet sein und ohne Förderung Rechtschreibschwierigkeiten entwickeln. Die Schüler der Nicht-Risikogruppe sollten keine Rechtschreibschwierigkeiten entwickeln. Bei der klassifikatorischen Analyse werden Prädiktoren in „Risiko“ und „Kein Risiko“ sowie Kriterien in „Betroffen“ und „Nicht betroffen“ dichotomisiert, sodass es vier mögliche Ergebnisse gibt: richtig und falsch positiv sowie richtig und falsch negativ. Für die Bestimmung der prognostischen Validität diagnostischer Verfahren sollten nachfolgende Kennwerte berechnet werden.

1. *Sensitivität*: Prozentualer Anteil der Problemschüler, die auch als Risikoschüler identifiziert wurden (richtig positiv).
2. *Spezifität*: Prozentualer Anteil der unauffälligen Schüler, die auch als Nicht-Risikoschüler identifiziert wurden (richtig negativ).
3. *Positiver prädiktiver Wert*: Wahrscheinlichkeit, mit der Risikoschüler tatsächlich Probleme entwickeln.
4. *Negativer prädiktiver Wert*: Wahrscheinlichkeit, mit der Nicht-Risikoschüler

tatsächlich keine Probleme entwickeln (P. Marx & Lenhard, 2011).

5. *Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote* (RATZ-Index): Der Index gibt an, „wie deutlich das Verfahren den Zufall „schlägt““ (P. Marx & Lenhard, 2011, S. 76). Ein Ratz-Index über 66 % deutet auf eine sehr gute, ein Index zwischen 34 % und 66 % auf eine gute, aber nicht spezifische Vorhersageleistung eines Verfahrens. Ein Index unter 34 % ist nicht akzeptabel (H. Marx, Jansen & Skowronek, 2000).
6. *Youden-Index* (Y-I): Der Index gibt an, wie gut ein diagnostisches Verfahren zwischen Fällen, die das Kriterium erfüllen („Betroffen“) und denen, die das Kriterium nicht erfüllen („Nicht betroffen“), differenzieren kann. Der Youden-Index wird wie folgt berechnet: $Y-I = \text{Sensitivität} + \text{Spezifität} - 1$. Der Youden-Index hat einen Wertebereich von $0 \leq Y-I \leq 1$. Je größer der Index ist, desto weniger falsch positive und falsch negative Ergebnisse gibt es und desto besser trennt das diagnostische Verfahren zwischen den genannten Fällen (Goldhammer & Hartig, 2007; Youden, 1950).
7. *Likelihood-Ratios* (LR): Wahrscheinlichkeitsverhältnisse, die das Verhältnis der Wahrscheinlichkeiten wiedergeben, dass bei Fällen, die das Kriterium erfüllen („Betroffen“) ein positives oder negatives Testergebnis eintritt, im Vergleich zur Wahrscheinlichkeit bei Fällen, die das Kriterium nicht erfüllen („Nicht-Betroffen“). Der positive LR+ gibt das Wahrscheinlichkeitsverhältnis eines positiven Testergebnisses bei Betroffenen und Nicht-Betroffenen wieder. Der positive LR+ wird wie folgt berechnet: $LR+ = \text{Sensitivität} / (1 - \text{Spezifität})$. Der negative LR- gibt das Wahrscheinlichkeitsverhältnis eines negativen Testergebnisses bei Betroffenen und Nicht-Betroffenen wieder. Der negative LR- wird wie folgt berechnet: $LR- = (1 - \text{Sensitivität}) / \text{Spezifität}$ (G. Schwarzer, Türp & Antes, 2002). In Anlehnung an Jaeschke, Guyatt und Sackett (1994) geben G. Schwarzer et al. (2002) Wertebereiche für die Wahrscheinlichkeitsverhältnisse der LR wieder. Die Wertebereiche der Wahrscheinlichkeitsverhältnisse und ihre Interpretation sind in Tabelle 13 aufgeführt.

Tabelle 13: Wertebereiche der Wahrscheinlichkeitsverhältnisse und ihre Interpretation (G. Schwarzer, Türp & Antes, 2002, S. 660)

Wahrscheinlichkeitsverhältnisse	Interpretation
LR+ > 10 bzw. LR- < 0,1	„überzeugende diagnostische Evidenz“
LR+ 5 - 10 bzw. LR- 0,1 - 0,2	„hohe diagnostische Evidenz“
LR + 2 - 5 bzw. LR- 0,2 - 0,5	„schwache diagnostische Evidenz“
LR+ 1 - 2 bzw. LR- 0,5 - 1	„kaum relevante diagnostische Evidenz“

Anmerkungen: LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio

5 Darstellung der Ergebnisse

In diesem Kapitel werden die Ergebnisse der Untersuchung dargestellt. Aus Gründen der Konsistenz erfolgt die Darstellung der Ergebnisse nach der Reihenfolge der Forschungsfragen und Hypothesen.

5.1 Stichprobenbeschreibung

Die Stichprobe umfasst Grundschüler der 2. Klassenstufen aus dem Landkreis Mecklenburg-Strelitz und der Stadt Neubrandenburg. Die Untersuchung fand an 13 Schulen, 10 Grundschulen und 3 Regionalen Schulen mit Grundschulteil, in 29 Klassen statt. Eine Klasse, die einzige Diagnoseförderklasse, die an der Untersuchung teilnahm, wurde im Nachhinein aus dem Datensatz entfernt, da deren Rechtschreibentwicklung aufgrund der Verteilung der Unterrichtsziele und -inhalte der ersten beiden Klassenstufen auf drei Schuljahre nicht mit der Rechtschreibentwicklung der übrigen Grundschulklassen vergleichbar ist. Insgesamt nahmen damit 425 Schüler mit vollständigen Datensätzen zu allen 4 Messzeitpunkten an der Untersuchung teil. Tabelle 14 gibt einen Überblick über die Verteilung der Stichprobe nach Region und Schultyp.

Tabelle 14: Verteilung der Stichprobe nach Region und Schultyp

Region	Schultyp		Gesamt N (%)
	GS	RSG	
Neubrandenburg N (%)	213 (50.1)	0	213 (50.1)
Mecklenburg-Strelitz N (%)	130 (30.6)	82 (19.3)	212 (49.9)
Gesamt N (%)	343 (80.7)	82 (19.3)	425

Anmerkungen: GS = Grundschule, RSG = Regionale Schule mit Grundschulteil, N = Anzahl

Die Stichprobe verteilte sich aufgrund des Untersuchungsdesigns auf zwei Teilstichproben, auf eine erste Gruppe und eine zweite Gruppe von Schulklassen. Die erste Gruppe führte die FE-RS 2 Form A/B in der Abfolge ABAB, die zweite Gruppe in der Abfolge BABA durch. Die erste Gruppe von Schulklassen umfasste 193 Schüler, die zweite Gruppe von Schulklassen 232 Schüler. Die Verteilung der Teilstichproben nach Region, Schultyp und Geschlecht ist in Tabelle 15 wiedergegeben.

Tabelle 15: Verteilung der Teilstichproben nach Region, Schultyp und Geschlecht

	Region		Schultyp		Geschlecht	
	NB	MST	GS	RSG	männl.	weibl.
1. Gruppe N (%)	108 (56.0)	85 (44.0)	172 (89.1)	21 (10.9)	97 (50.3)	96 (49.7)
2. Gruppe N (%)	105 (45.3)	127 (54.7)	171 (73.7)	61 (26.3)	105 (45.3)	127 (54.7)

Anmerkungen: NB = Neubrandenburg, MST = Mecklenburg-Strelitz, GS = Grundschule, RSG = Regionale Schule mit Grundschulteil, männl. = männlich, weibl. = weiblich, N = Anzahl

Die Gesamtstichprobe von 425 Schülern verteilte sich auf 202 Jungen (47.5 %) und 223 Mädchen (52.5 %). Tabelle 16 gibt einen Überblick über die Altersverteilung der Gesamtstichprobe zu den 4 Messzeitpunkten.

Tabelle 16: Altersverteilung der Stichprobe zu den 4 Messzeitpunkten

	M (in Jahren)	SD (in Monaten)	Min. (in Jahren)	Max. (in Jahren)
1. MZP	7.5	4.8	6	10
2. MZP	7.7	4.8	7	10
3. MZP	7.8	4.8	7	10
4. MZP	8.0	4.8	7	10

Anmerkungen: MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, Min. = Minimum, Max. = Maximum

5.2 Darstellung der Ergebnisse der FE-RS 2

Im Folgenden werden die Ergebnisse der FE-RS 2 dargelegt. Entsprechend der Reihenfolge der Forschungsfragen und Hypothesen werden Ergebnisse zur Differenzierungsfähigkeit, Reliabilität, Validität, prognostischen Validität, Äquivalenz der Form A und Form B sowie zur Veränderungssensitivität des Verfahrens dargestellt.

5.2.1 Differenzierungsfähigkeit

In diesem Abschnitt werden zuerst die Häufigkeitsverteilungen sowie Mittelwerte und Standardabweichungen für den orthographischen Gesamtwert und die Subskalen der FE-RS 2 über alle 4 Messzeitpunkte dargelegt. Anschließend folgen die Ergebnisse der Analysen zu Schwierigkeiten und Trennschärfen.

5.2.1.1 Häufigkeitsverteilungen, Mittelwerte und Standardabweichungen der Form

A

Die Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form A über alle 4 Messzeitpunkte sind in Abbildung 18 dargestellt.

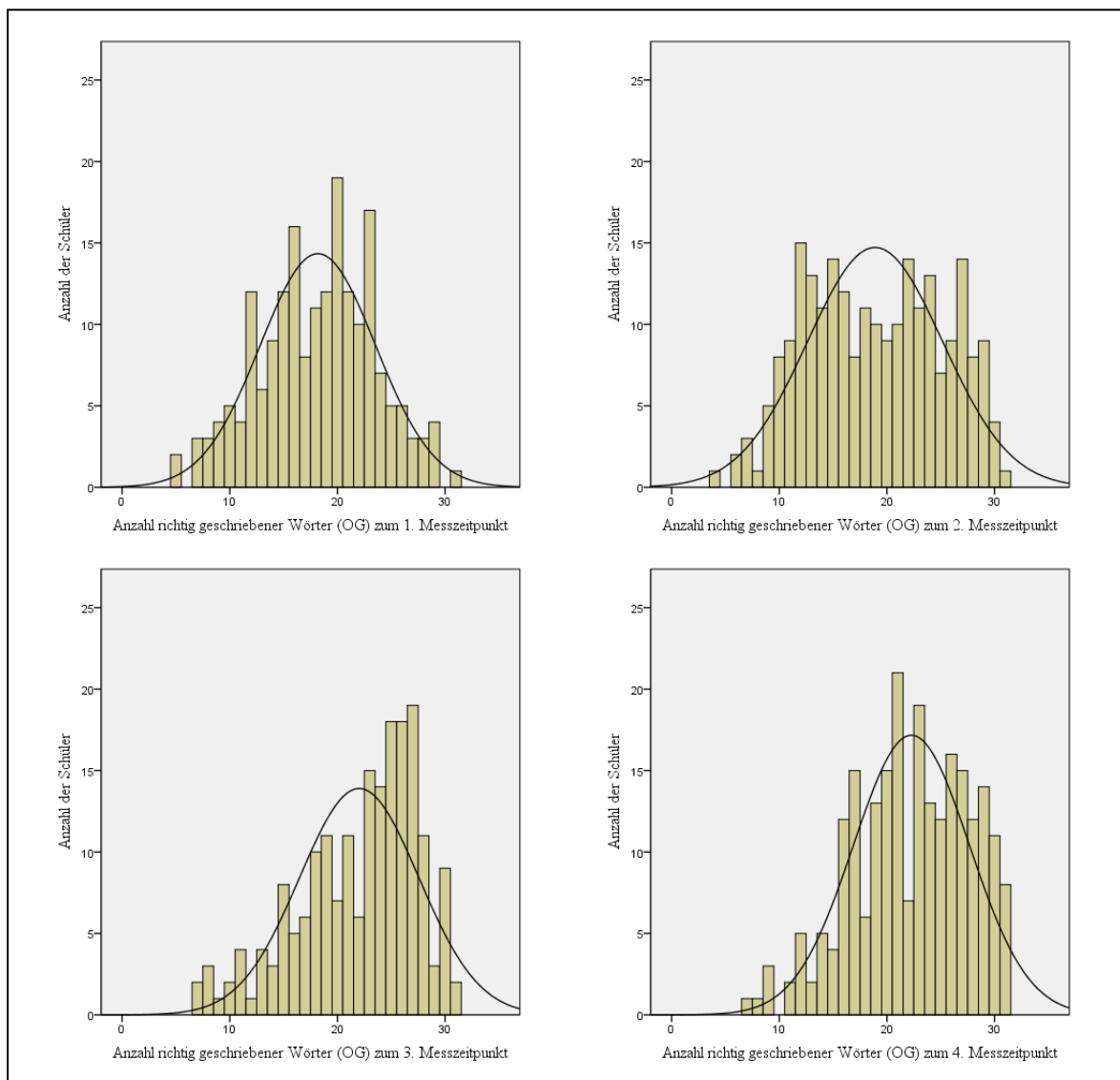


Abbildung 18: Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form A

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, OG = orthographischer Gesamtwert, N = 193 für Form A zum 1. und 3. Messzeitpunkt, N = 232 für Form A zum 2. und 4. Messzeitpunkt

Wie in Abbildung 18 zu erkennen, ist die Verteilungsform für den orthographischen Gesamtwert der Form A zum 1. Messzeitpunkt symmetrisch, eingipflig und glockenförmig. Nach der Prüfung mit dem KST entspricht die Verteilung einer

Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 5$ im Mittel 18 von 31 Wörtern richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für den orthographischen Gesamtwert der Form A symmetrisch mit der Tendenz zur Zweigipfligkeit, die sich einer breitgipfligen Verteilung annähert. Nach der Prüfung mit dem KST entspricht die Verteilung dennoch einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 6$ im Mittel 19 von 31 Wörtern richtig geschrieben. Zum 3. Messzeitpunkt ist die Verteilungsform für den orthographischen Gesamtwert der Form A asymmetrisch, eingipflig und linksschief. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 5$ im Mittel 22 von 31 Wörtern richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für den orthographischen Gesamtwert der Form A asymmetrisch, eingipflig, leicht linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung dennoch einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 5$ im Mittel 22 von 31 Wörtern richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für den orthographischen Gesamtwert Verschiebungen von Normalverteilungen zu einer linksschiefen Verteilung. Das heißt, der orthographische Gesamtwert der Form A differenziert zu den ersten drei Messzeitpunkten gut zwischen Zweitklässlern mit guten und schlechten Rechtschreibleistungen und für den 3. Messzeitpunkt ermöglicht die Form A eine gute Differenzierung im unteren Leistungsbereich. Zudem zeigen die Verschiebungen, dass die Schüler von Messzeitpunkt zu Messzeitpunkt mehr Wörter orthographisch richtig schrieben. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 17 aufgeführt sind, groß sind. Dies bestätigt, dass der orthographische Gesamtwert der Form A zwischen Zweitklässlern mit guten und schlechten Rechtschreibleistungen differenzieren kann. Tabelle J1 im Anhang J gibt die Ergebnisse der Überprüfung auf Normalverteilung des orthographischen Gesamtwerts der Form A mit dem KST wieder. Tabelle 17 gibt die Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für den orthographischen Gesamtwert der Form A über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 17: Kennwerte des orthographischen Gesamtwerts der Form A über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	193	18.2	5.4	2.2	5	31
2. MZP	232	18.9	6.3	2.3	4	31
3. MZP	193	22.0	5.5	2.1	7	31
4. MZP	232	22.3	5.4	2.1	7	31

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphakoeffizient, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilungen für die ELG-Subskala der Form A über alle 4 Messzeitpunkte sind in Abbildung 19 dargestellt.

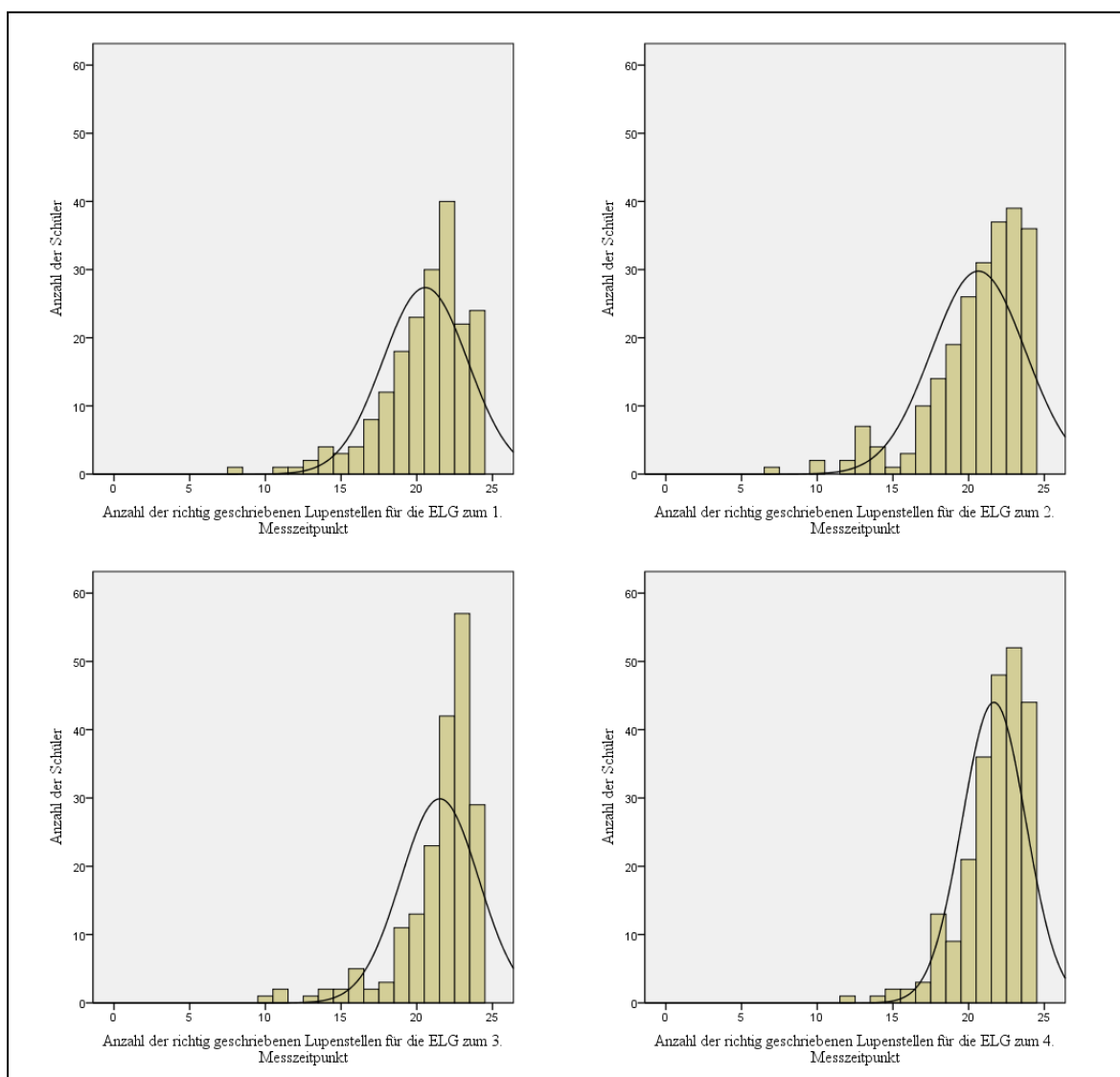


Abbildung 19: Häufigkeitsverteilungen für die ELG-Subskala der Form A

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELG = elementar graphematisches Können, N = 193 für Form A zum 1. und 3. Messzeitpunkt, N = 232 für Form A zum 2. und 4. Messzeitpunkt

Wie in Abbildung 19 zu erkennen, ist die Verteilungsform für die ELG-Subskala der Form A zum 1. Messzeitpunkt asymmetrisch, eingipflig, linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 21 von 24 ELG-Lupenstellen richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für die ELG-Subskala der Form A asymmetrisch, eingipflig, linksschief, tendenziell schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 21 von 24 ELG-Lupenstellen richtig geschrieben. Zum 3. Messzeitpunkt ist die Verteilungsform für die ELG-Subskala der Form A asymmetrisch, eingipflig, linksschief, sehr schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 21 von 24 ELG-Lupenstellen richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für die ELG-Subskala der Form A asymmetrisch, eingipflig, linksschief, sehr schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 22 von 24 ELG-Lupenstellen richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für die ELG-Subskalen Verschiebungen von linksschiefen zu schmalgipfligen, linksschiefen Verteilungen. Das heißt, die ELG-Subskala der Form A ermöglicht eine gute Differenzierung im unteren Leistungsbereich. Zudem zeigen die schmalgipfligen, linksschiefen Verteilungen, dass die Mehrzahl der Schüler zum Ende der 2. Klassenstufe die ELG-Lupenstellen richtig schrieben. Des Weiteren zeigen sich über alle 4 Messzeitpunkte kaum Veränderungen in den Mittelwerten. Darüber hinaus zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 18 aufgeführt sind, nicht sehr groß sind. Das heißt, dass sich mit der ELG-Subskala der Form A Schüler nicht in mehrere Gruppen differenzieren lassen. Tabelle J2 im Anhang J gibt die Ergebnisse der Überprüfung auf Normalverteilung der ELG-Subskala der Form A mit dem KST wieder. Tabelle 18 gibt die Mittelwerte, Standardabweichungen, Standardmessfehler, das Minimum und Maximum für die ELG-Subskala der Form A über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 18: Kennwerte der ELG-Subskala der Form A über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	193	20.6	2.8	1.5	8	24
2. MZP	232	20.6	3.1	1.5	7	24
3. MZP	193	21.5	2.6	1.4	10	24
4. MZP	232	21.7	2.1	1.3	12	24

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, ELG = elementar graphematisches Können, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphaeffizient, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilungen für die ELO-Subskala der Form A über alle 4 Messzeitpunkte sind in Abbildung 20 dargestellt.

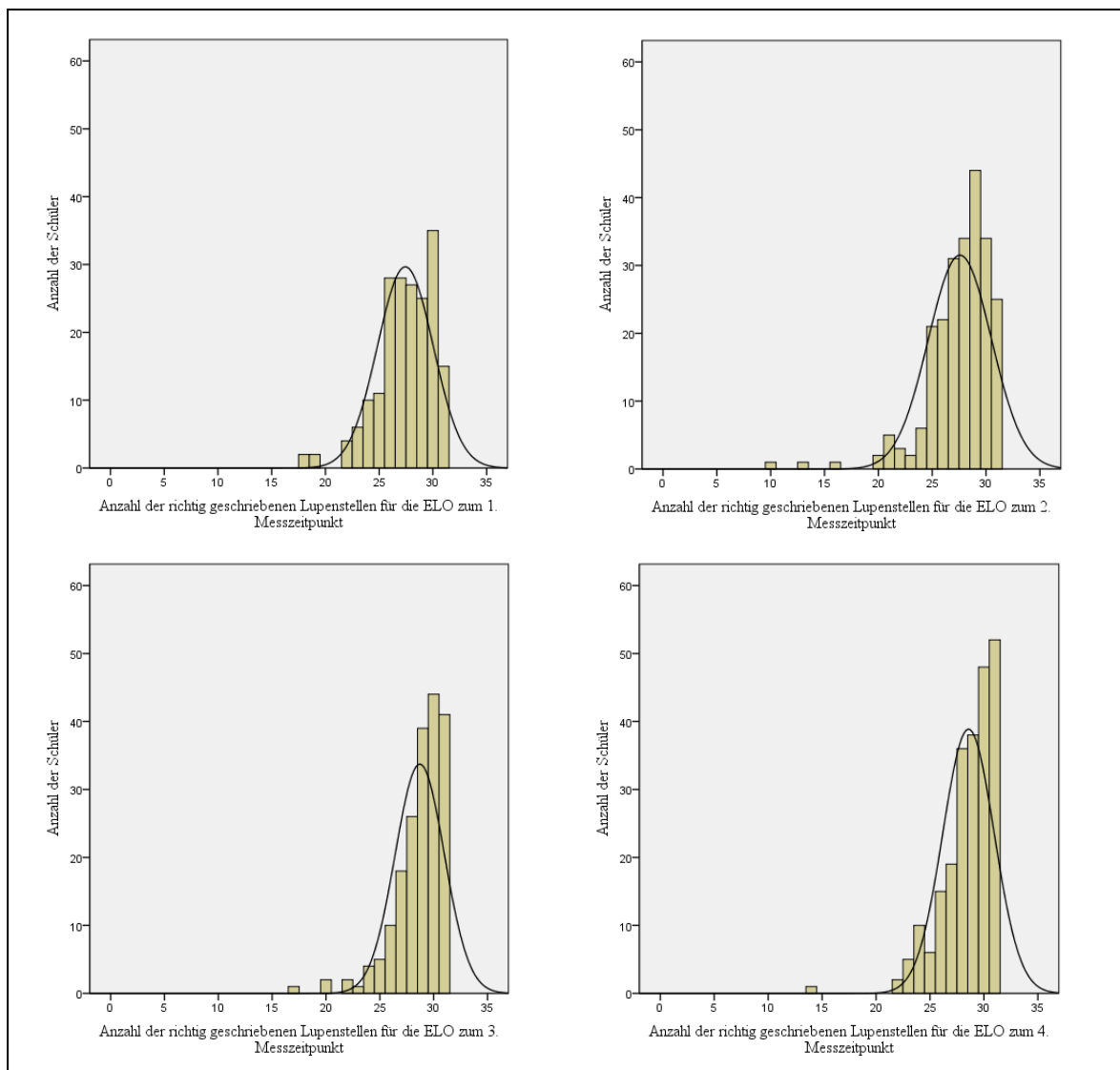


Abbildung 20: Häufigkeitsverteilungen für die ELO-Subskala der Form A

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELO = elementar orthographisches Können, N = 193 für Form A zum 1. und 3. Messzeitpunkt, N = 232 für Form A zum 2. und 4. Messzeitpunkt

Wie in Abbildung 20 zu erkennen, ist die Verteilungsform für die ELO-Subskala der Form A zum 1. Messzeitpunkt asymmetrisch, eingipflig, leicht linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 27 von 31 ELO-Lupenstellen richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für die ELO-Subskala der Form A leicht asymmetrisch, eingipflig, leicht linksschief, tendenziell schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 28 von 31 ELO-Lupenstellen richtig geschrieben. Zum 3. Messzeitpunkt ist die Verteilungsform für die ELO-Subskala der Form A asymmetrisch, eingipflig, linksschief, sehr schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 29 von 31 ELO-Lupenstellen richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für die ELO-Subskala der Form A asymmetrisch, eingipflig, linksschief, sehr schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 29 von 31 ELO-Lupenstellen richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für die ELO-Subskalen Verschiebungen von linksschiefen zu schmalgipfligen, linksschiefen Verteilungen. Das heißt, die ELO-Subskala der Form A ermöglicht eine gute Differenzierung im unteren Leistungsbereich. Zudem zeigen die schmalgipfligen, linksschiefen Verteilungen, dass die Mehrzahl der Schüler zum Ende der 2. Klassenstufe die ELO-Lupenstellen richtig schrieben. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 19 aufgeführt sind, nicht sehr groß sind. Das heißt, dass sich mit der ELO-Subskala der Form A Schüler nicht in mehrere Gruppen differenzieren lassen. Tabelle J3 im Anhang J gibt die Ergebnisse der Überprüfung auf Normalverteilung der ELO-Subskala der Form A mit dem KST wieder. Tabelle 19 gibt die Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für die ELO-Subskala der Form A über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 19: Kennwerte der ELO-Subskala der Form A über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	193	27.4	2.6	1.5	18	31
2. MZP	232	27.6	2.9	1.5	10	31
3. MZP	193	28.7	2.3	1.3	17	31
4. MZP	232	28.6	2.4	1.4	14	31

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, ELO = elementar orthographisches Können, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alpkoeffizient, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilungen für die ERG-Subskala der Form A über alle 4 Messzeitpunkte sind in Abbildung 21 dargestellt.

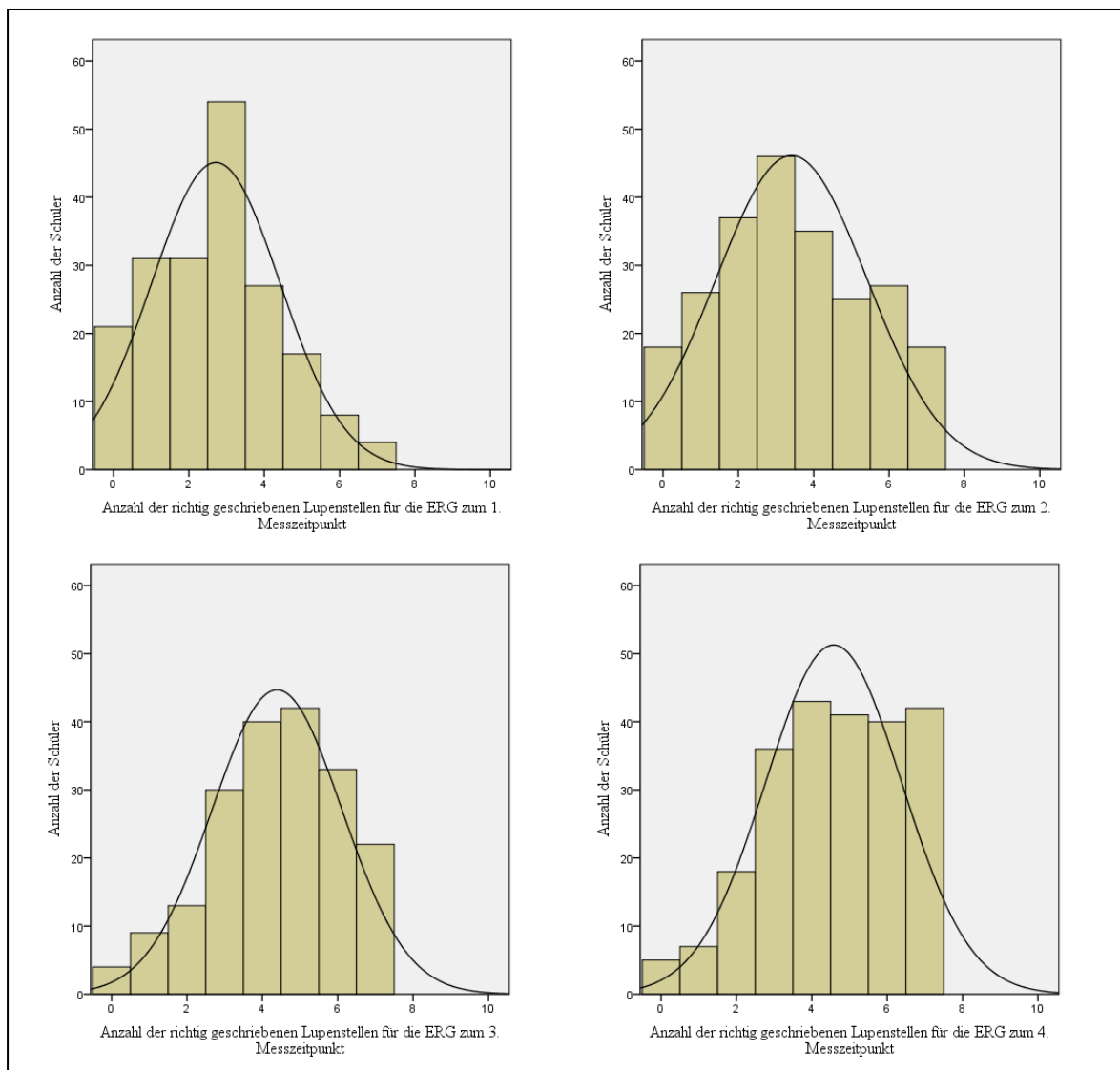


Abbildung 21: Häufigkeitsverteilungen für die ERG-Subskala der Form A

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ERG = erweitert graphematisches Können, N = 193 für Form A zum 1. und 3. Messzeitpunkt, N = 232 für Form A zum 2. und 4. Messzeitpunkt

Wie in Abbildung 21 zu erkennen, ist die Verteilungsform für die ERG-Subskala der Form A zum 1. Messzeitpunkt asymmetrisch, eingipflig, tendenziell rechtsschief und der Auslauf der Verteilungsform ist links abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 3 von 7 ERG-Lupenstellen richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für die ERG-Subskala der Form A annähernd symmetrisch, eingipflig, jedoch nicht glockenförmig. Der Auslauf der Verteilungsform ist links abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 3 von 7 ERG-Lupenstellen richtig geschrieben. Zum 3. Messzeitpunkt ist die Verteilungsform für die ERG-Subskala der Form A asymmetrisch, eingipflig, tendenziell linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 4 von 7 ERG-Lupenstellen richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für die ERG-Subskala der Form A asymmetrisch, eingipflig, leicht linksschief, tendenziell breitgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 5 von 7 ERG-Lupenstellen richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für die ERG-Subskalen Verschiebungen von rechtsschiefen zu linksschiefen Verteilungen. Das heißt, die ERG-Subskala der Form A ermöglicht zum 1. Messzeitpunkt eine gute Differenzierung im oberen und zum 3. Messzeitpunkt im unteren Leistungsbereich. Zudem zeigen die Verschiebungen, dass die Schüler von Messzeitpunkt zu Messzeitpunkt mehr ERG-Lupenstellen richtig schrieben und dass die Schüler die Vokalquantität später als die elementare Stufe, wie z. B. alphabetische Schreibungen, Groß- und Kleinschreibung oder Auslautverhärtungen, beherrschten. Letzteres entspricht dem aktuellen Forschungsstand der Rechtschreibentwicklung und dem Vorgehen im Rechtschreibunterricht. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 20 aufgeführt sind, nicht sehr groß sind. Das heißt, dass sich mit der ERG-Subskala der Form A Schüler nicht in mehrere Gruppen differenzieren lassen. Tabelle J4 im Anhang J gibt die Ergebnisse der Überprüfung auf Normalverteilung der ERG-Subskala der Form A mit dem KST wieder. Tabelle 20 gibt die

Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für die ERG-Subskala der Form A über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 20: Kennwerte der ERG-Subskala der Form A über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	193	2.7	1.7	1.0	0	7
2. MZP	232	3.4	2.0	1.0	0	7
3. MZP	193	4.4	1.7	1.0	0	7
4. MZP	232	4.6	1.8	1.0	0	7

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphakoeffizient, Min. = Minimum, Max. = Maximum

5.2.1.2 Häufigkeitsverteilungen, Mittelwerte und Standardabweichungen der Form B

Die Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form B über alle 4 Messzeitpunkte sind in Abbildung 22 dargestellt.

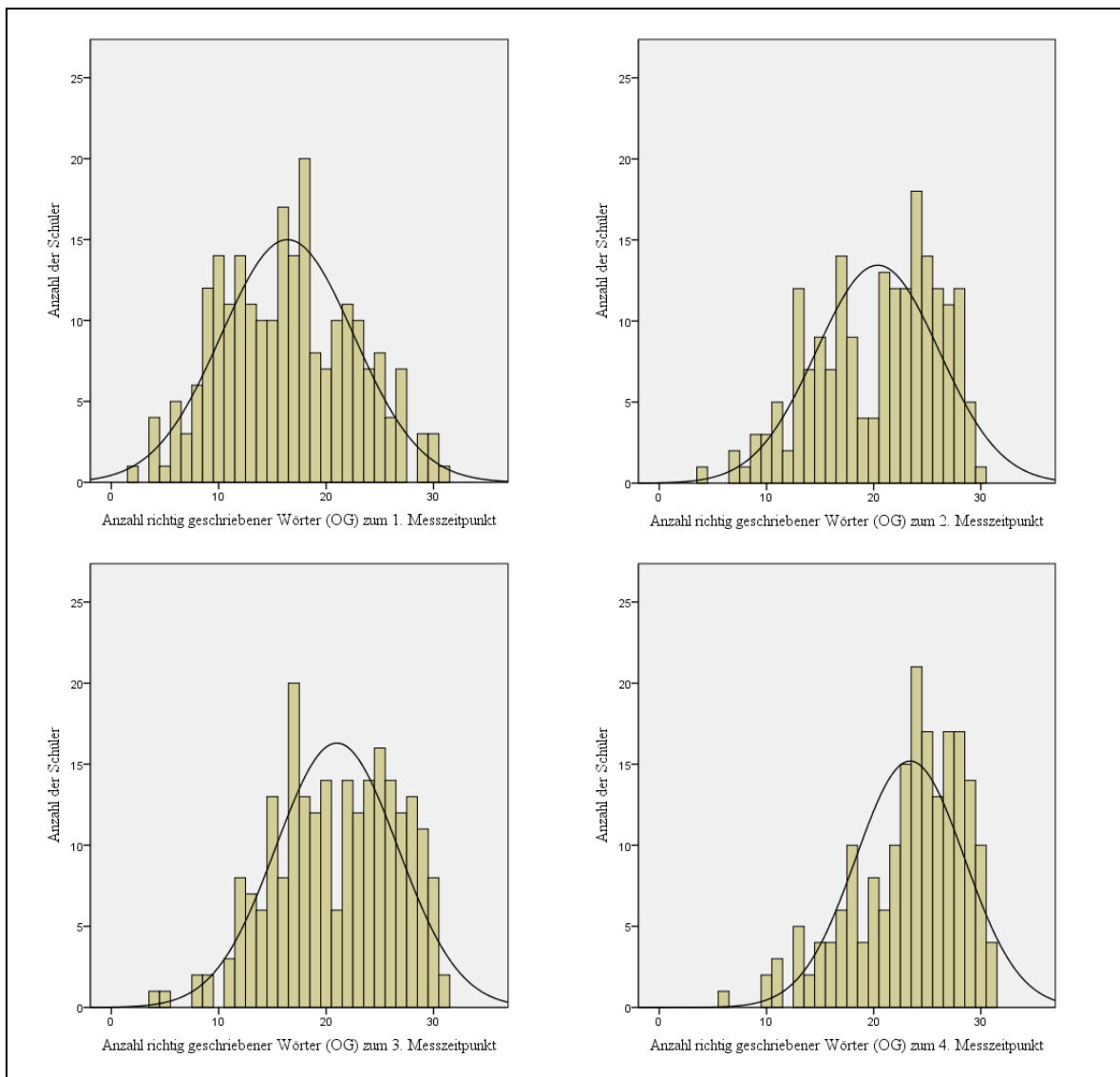


Abbildung 22: Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form B

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, N = 232 für Form B zum 1. und 3. Messzeitpunkt, N = 193 für Form B zum 2. und 4. Messzeitpunkt

Wie in Abbildung 22 zu erkennen, ist die Verteilungsform für den orthographischen Gesamtwert der Form B zum 1. Messzeitpunkt annähernd symmetrisch, eingipflig und annähernd glockenförmig. Nach der Prüfung mit dem KST entspricht die Verteilung einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 6$ im Mittel 16 von 31 Wörtern richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für den orthographischen Gesamtwert der Form B asymmetrisch, zweigipflig und tendenziell linksschief. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 6$ im Mittel 20 von 31 Wörtern richtig geschrieben. Zum 3.

Messzeitpunkt ist die Verteilungsform für den orthographischen Gesamtwert der Form B annähernd symmetrisch, eingipflig und sehr leicht linksschief. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 6$ im Mittel 21 von 31 Wörtern richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für den orthographischen Gesamtwert der Form B asymmetrisch, eingipflig und linksschief. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 5$ im Mittel 28 von 31 Wörtern richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für den orthographischen Gesamtwert Verschiebungen von einer Normalverteilung zu linksschiefen Verteilungen. Das heißt, der orthographische Gesamtwert der Form B ermöglicht eine gute Differenzierung im unteren Leistungsbereich. Zudem zeigen die Verschiebungen, dass die Schüler von Messzeitpunkt zu Messzeitpunkt mehr Wörter orthographisch richtig schrieben. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 22 aufgeführt sind, groß sind. Dies bestätigt, dass der orthographische Gesamtwert der Form B zwischen Zweitklässlern mit guten und schlechten Rechtschreibleistungen differenzieren kann. Tabelle K1 im Anhang K gibt die Ergebnisse der Überprüfung auf Normalverteilung des orthographischen Gesamtwerts der Form B mit dem KST wieder. Tabelle 21 gibt die Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für den orthographischen Gesamtwert der Form B über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 21: Kennwerte des orthographischen Gesamtwerts der Form B über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	232	16.4	6.2	2.1	2	31
2. MZP	193	20.4	5.7	2.1	4	30
3. MZP	232	21.0	5.7	2.1	4	31
4. MZP	193	23.4	5.1	2.0	6	31

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphakoeffizient, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilungen für die ELG-Subskala der Form B über alle 4 Messzeitpunkte sind in Abbildung 23 dargestellt.

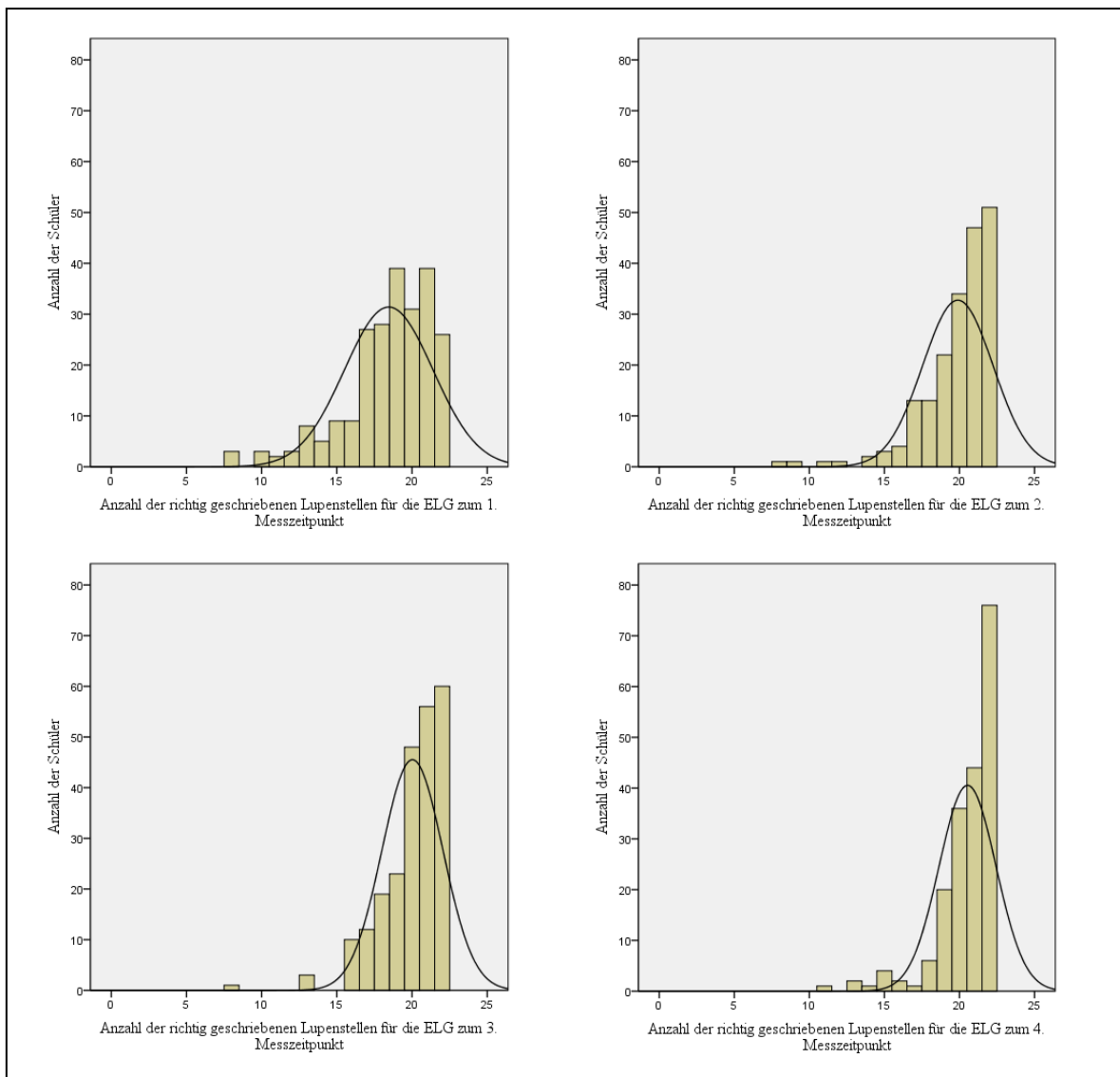


Abbildung 23: Häufigkeitsverteilungen für die ELG-Subkala der Form B

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELG = elementar graphematisches Können, N = 232 für Form B zum 1. und 3. Messzeitpunkt, N = 193 für Form B zum 2. und 4. Messzeitpunkt

Wie in Abbildung 23 zu erkennen, ist die Verteilungsform für die ELG-Subkala der Form B zum 1. Messzeitpunkt asymmetrisch, eingipflig, linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 18 von 22 ELG-Lupenstellen richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für die ELG-Subkala der Form B asymmetrisch, eingipflig, linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 20 von 22 ELG-Lupenstellen richtig geschrieben. Zum 3. Messzeitpunkt

ist die Verteilungsform für die ELG-Subskala der Form B asymmetrisch, eingipflig, linksschief, tendenziell schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 20 von 22 ELG-Lupenstellen richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für die ELG-Subskala der Form B asymmetrisch, eingipflig, linksschief, sehr schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 20 von 22 ELG-Lupenstellen richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für die ELG-Skalen Verschiebungen von linksschiefen zu schmalgipfligen, linksschiefen Verteilungen. Das heißt, die ELG-Subskala der Form B ermöglicht eine gute Differenzierung im unteren Leistungsbereich. Zudem zeigen die schmalgipfligen, linksschiefen Verteilungen, dass die Mehrzahl der Schüler zum Ende der 2. Klassenstufe die ELG-Lupenstellen richtig schrieben. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 23 aufgeführt sind, nicht sehr groß sind. Das heißt, dass sich mit der ELG-Subskala der Form B Schüler nicht in mehrere Gruppen, z. B. Schüler mit sehr schlechten, schlechten, guten oder sehr guten Rechtschreibleistungen, differenzieren lassen. Tabelle K2 im Anhang K gibt die Ergebnisse der Überprüfung auf Normalverteilung der ELG-Subskala der Form B mit dem KST wieder. Tabelle 22 gibt die Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für die ELG-Subskala der Form B über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 22: Kennwerte der ELG-Subskala der Form B über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	232	18.5	2.9	1.5	8	22
2. MZP	193	19.9	2.3	1.2	8	22
3. MZP	232	20.0	2.0	1.2	8	22
4. MZP	193	20.5	1.9	1.1	11	22

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, ELG = elementar graphematisches Können, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphakoeffizient, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilungen für die ELO-Subskala der Form B über alle 4 Messzeitpunkte sind in Abbildung 24 dargestellt.

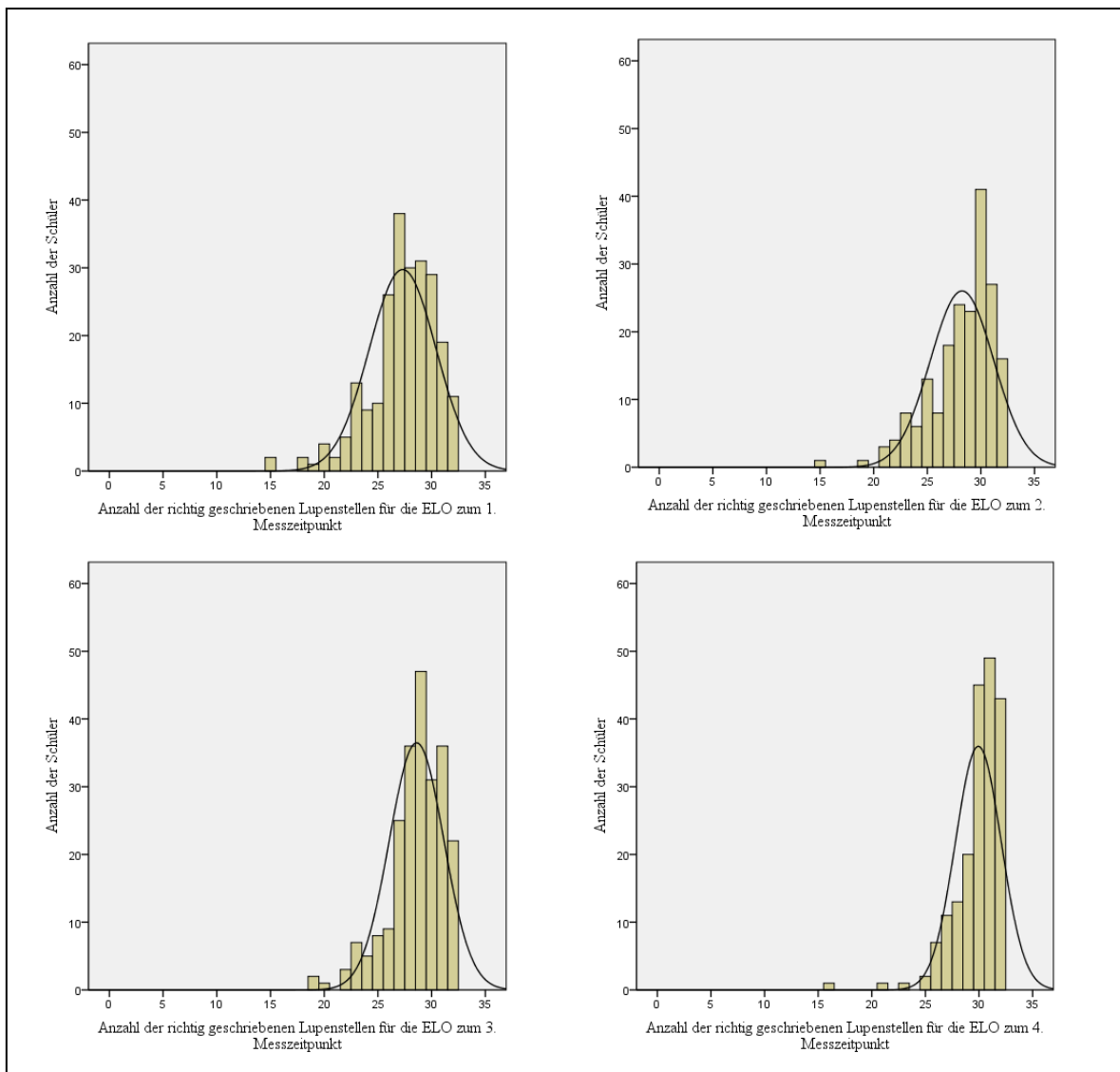


Abbildung 24: Häufigkeitsverteilungen für die ELO-Subskala der Form B
Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELO = elementar orthographisches Können, N = 232 für Form B zum 1. und 3. Messzeitpunkt, N = 193 für Form B zum 2. und 4. Messzeitpunkt

Wie in Abbildung 24 zu erkennen, ist die Verteilungsform für die ELO-Subskala der Form B zum 1. Messzeitpunkt asymmetrisch, eingipflig und leicht linksschief. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 27 von 32 ELO-Lupenstellen richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für die ELO-Subskala der Form B asymmetrisch, eingipflig, linksschief und der Auslauf der Verteilungsform ist tendenziell rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 3$ im Mittel 28 von 32 ELO-Lupenstellen richtig geschrieben. Zum 3. Messzeitpunkt ist die Verteilungsform für die ELO-Subskala der

Form B asymmetrisch, eingipflig, linksschief, tendenziell schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 29 von 32 ELO-Lupenstellen richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für die ELO-Subskala der Form B asymmetrisch, eingipflig, linksschief, schmalgipflig und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 30 von 32 ELO-Lupenstellen richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für die ELO-Subskalen Verschiebungen von linksschiefen zu schmalgipfligen, linksschiefen Verteilungen. Das heißt, die ELO-Subskala der Form B ermöglicht eine gute Differenzierung im unteren Leistungsbereich. Zudem zeigen die schmalgipfligen, linksschiefen Verteilungen, dass die Mehrzahl der Schüler zum Ende der 2. Klassenstufe die ELO-Lupenstellen richtig schrieben. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 24 aufgeführt sind, nicht sehr groß sind. Das heißt, dass sich mit der ELO-Subskala der Form B Schüler nicht in mehrere Gruppen differenzieren lassen. Tabelle K3 im Anhang K gibt die Ergebnisse der Überprüfung auf Normalverteilung der ELO-Subskala der Form B mit dem KST wieder. Tabelle 23 gibt die Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für die ELO-Subskala der Form B über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 23: Kennwerte der ELO-Subskala der Form B über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	232	27.3	3.1	1.6	15	32
2. MZP	193	28.2	2.9	1.5	15	32
3. MZP	232	28.6	2.5	1.4	19	32
4. MZP	193	29.9	2.1	1.2	16	32

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, ELO = elementar orthographisches Können, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphakoeffizient, Min. = Minimum, Max. = Maximum

Die Häufigkeitsverteilungen für die ERG-Subskala der Form B über alle 4 Messzeitpunkte sind in Abbildung 25 dargestellt.

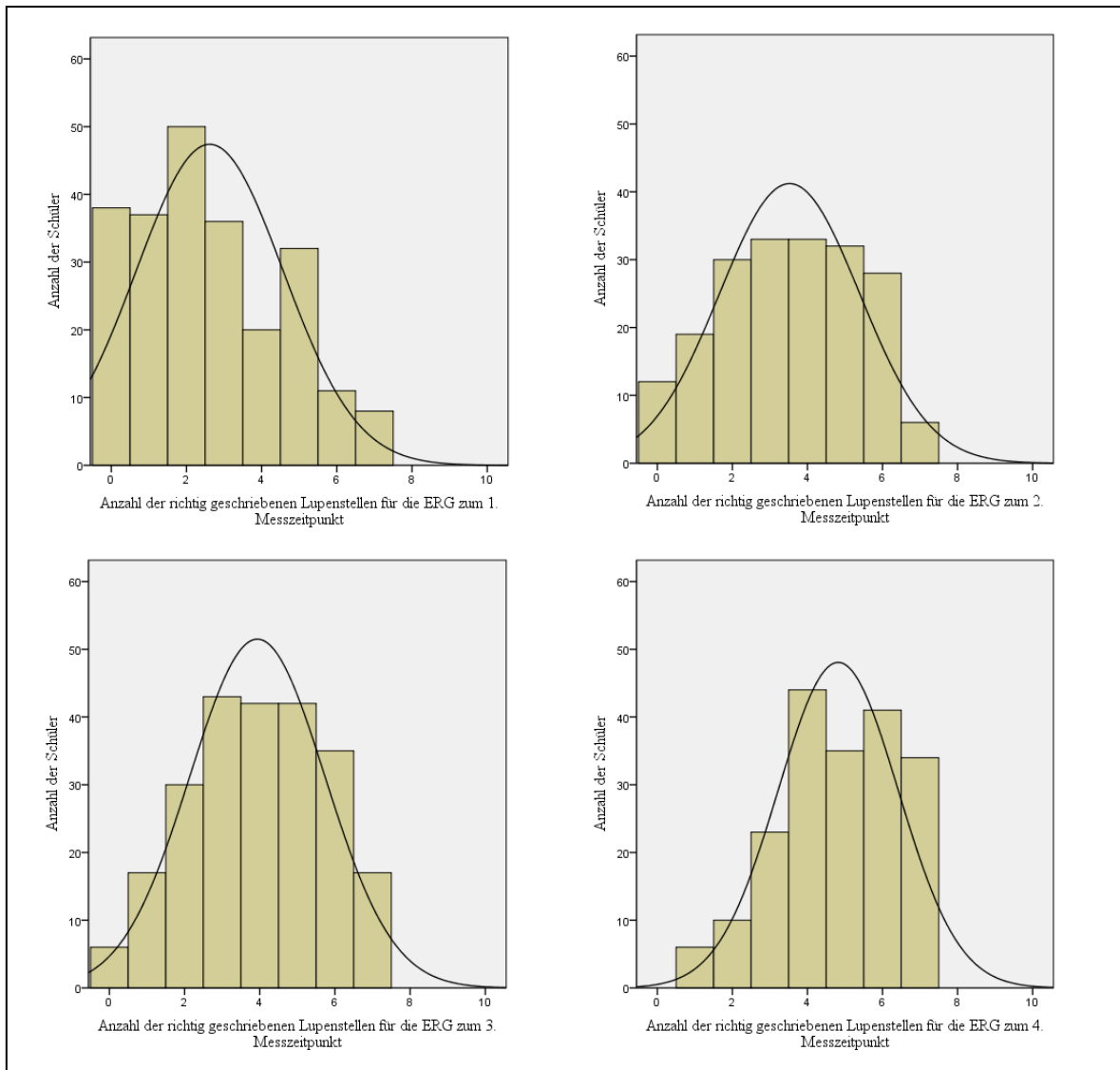


Abbildung 25: Häufigkeitsverteilungen für die ERG-Subskala der Form B
Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ERG = erweitert graphematisches Können, N = 232 für Form B zum 1. und 3. Messzeitpunkt, N = 193 für Form B zum 2. und 4. Messzeitpunkt

Wie in Abbildung 25 zu erkennen, ist die Verteilungsform für die ERG-Subskala der Form B zum 1. Messzeitpunkt asymmetrisch, eingipflig, rechtsschief und der Auslauf der Verteilungsform ist links abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 1. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 3 von 7 ERG-Lupenstellen richtig geschrieben. Zum 2. Messzeitpunkt ist die Verteilungsform für die ERG-Subskala der Form B annähernd symmetrisch, eingipflig, jedoch nicht glockenförmig. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 2. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 3 von 7 ERG-Lupenstellen richtig geschrieben. Zum 3. Messzeitpunkt ist die Verteilungsform für die ERG-Subskala

der Form B annähernd symmetrisch, eingipflig und annähernd glockenförmig. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 3. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 4 von 7 ERG-Lupenstellen richtig geschrieben. Zum 4. Messzeitpunkt ist die Verteilungsform für die ERG-Subskala der Form B asymmetrisch, eingipflig, linksschief und der Auslauf der Verteilungsform ist rechts abgebrochen. Nach der Prüfung mit dem KST entspricht die Verteilung nicht einer Normalverteilung. Zum 4. Messzeitpunkt wurden bei einer Standardabweichung von $SD = 2$ im Mittel 5 von 7 ERG-Lupenstellen richtig geschrieben. Zusammenfassend zeigen sich über alle 4 Messzeitpunkte für die ERG-Subskalen Verschiebungen von rechtsschiefen zu linksschiefen Verteilungen. Das heißt, die ERG-Subskala der Form B ermöglicht zum 1. Messzeitpunkt eine gute Differenzierung im oberen und zum 4. Messzeitpunkt im unteren Leistungsbereich. Zudem zeigen die Verschiebungen, dass die Schüler von Messzeitpunkt zu Messzeitpunkt mehr ERG-Lupenstellen richtig schrieben und dass die Schüler die Vokalquantität später als die elementaren Strategien beherrschten. Des Weiteren zeigt sich über alle 4 Messzeitpunkte, dass die Standardabweichungen im Vergleich zu den Standardmessfehlern, die in Tabelle 25 aufgeführt sind, nicht sehr groß sind. Das heißt, dass sich mit der ERG-Subskala der Form B Schüler nicht in mehrere Gruppen differenzieren lassen. Tabelle K4 im Anhang K gibt die Ergebnisse der Überprüfung auf Normalverteilung der ERG-Subskala der Form B mit dem KST wieder. Tabelle 24 gibt die Mittelwerte, Standardabweichungen und Standardmessfehler sowie das Minimum und Maximum für die ERG-Subskala der Form B über alle 4 Messzeitpunkte zusammenfassend wieder.

Tabelle 24: Kennwerte der ERG-Subskala der Form B über alle 4 Messzeitpunkte

	N	M	SD	SE	Min.	Max.
1. MZP	232	2.6	1.9	1.0	0	7
2. MZP	193	3.5	1.9	1.1	0	7
3. MZP	232	3.9	1.8	1.1	0	7
4. MZP	193	4.8	1.6	1.0	1	7

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, M = Mittelwert, SD = Standardabweichung, SE = Standardmessfehler berechnet mit Alphakoeffizient, Min. = Minimum, Max. = Maximum

H_{1,1}: Die Häufigkeitsverteilungen für den orthographischen Gesamtwert und die Subskalen der Form A und der Form B sind für alle 4 Messzeitpunkte glockenförmig, symmetrisch und eingipflig.

Zusammenfassend betrachtet, sind die Häufigkeitsverteilungen für den orthographischen Gesamtwert und die Subskalen der FE-RS 2 nicht über alle 4 Messzeitpunkte normalverteilt. Nur für den orthographischen Gesamtwert der Form A des 1., 2. und 4. Messzeitpunkts und für den orthographischen Gesamtwert der Form B des 1. Messzeitpunkts sind die Daten normalverteilt. Alle übrigen Daten des orthographischen Gesamtwerts und der Subskalen sind fast ausnahmslos linksschief verteilt, es traten dort Deckeneffekte auf. Die Hypothese $H_{1,1}$ lässt sich deshalb nur zum Teil bestätigen. Die FE-RS 2 differenziert überwiegend gut im unteren und mittleren Leistungsbereich. Schüler mit sehr guten Rechtschreibleistungen werden durch die FE-RS 2 zum Teil nicht differenziert erfasst. Für die Anwendung der FE-RS 2 ist Letzteres aber keine Einschränkung, da mit der FE-RS 2 Schüler mit Rechtschreibschwierigkeiten, das heißt Schüler im unteren Leistungsbereich, frühzeitig identifiziert und deren Rechtschreibschwierigkeiten qualitativ eingeschätzt werden sollen.

$H_{1,2}$: Die Standardabweichungen der Messwerte für den orthographischen Gesamtwert und die Subskalen der Form A und der Form B sind für alle 4 Messzeitpunkte hoch.

Die Standardabweichungen für den orthographischen Gesamtwert und die Subskalen der FE-RS 2 sind nicht über alle 4 Messzeitpunkte hoch. Nur für den orthographischen Gesamtwert und die ERG-Subskala sind die Standardabweichungen hoch. Die Hypothese $H_{1,2}$ lässt sich deshalb nur zum Teil bestätigen. Die Folge ist, dass (bei hoher Reliabilität [Forschungsfrage 2]) nur der orthographische Gesamtwert der FE-RS 2 dazu geeignet ist, Schüler mit unterschiedlichen Rechtschreibleistungen in verschiedene Gruppen zu differenzieren. Die Subskalen der FE-RS 2 sind, bis auf die ERG-Subskala, dazu nicht in der Lage.

5.2.1.3 Itemschwierigkeiten und Trennschärfen

Für die weitere Beantwortung der Forschungsfrage zur Differenzierungsfähigkeit des Testverfahrens wurden die Itemschwierigkeiten und Trennschärfen der FE-RS 2 berechnet. Die Berechnung erfolgte für den orthographischen Gesamtwert und für die Subskalen über alle 4 Messzeitpunkte.

Beide Formen der FE-RS 2 decken Items mit unterschiedlichen Schwierigkeitsgraden und Trennschärfen ab. Die Schwierigkeitsindexe steigen für den orthographischen Gesamtwert und die Subskalen innerhalb der FE-RS 2 an. Der Schwierigkeitsgrad der zu schreibenden Wörter nimmt also ab. Die Trennschärfen der einzelnen Items liegen im niedrigen bis hohen Bereich. Items mit einem hohen Schwierigkeitsindex, z. B. *Oma*, *Ofen*, *Name* oder *Baum*, haben eine niedrige Trennschärfe und differenzieren zwischen Schülern mit unterschiedlichen Rechtschreibleistungen nur ungenügend. Alle einzelnen Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form A über alle 4 Messzeitpunkte befinden sich in den Tabellen L1 bis L5 im Anhang L, für die Form B in den Tabellen M1 bis M5 im Anhang M.

Die gemittelten Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form A über alle 4 Messzeitpunkte gibt Tabelle 25 zusammenfassend wieder.

Tabelle 25: Gemittelte Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form A über alle 4 Messzeitpunkte

	Gemittelte Itemschwierigkeit p_i				Gemittelte Trennschärfe r_{it}			
	1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
OG	.59	.61	.71	.72	.33	.39	.36	.35
ELG	.86	.86	.89	.90	.25	.32	.28	.20
ELO	.88	.89	.92	.92	.21	.26	.25	.24
ERG	.39	.49	.62	.65	.38	.46	.34	.39

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, ERO = erweitert orthographisches Können

Die gemittelten Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form B über alle 4 Messzeitpunkte gibt Tabelle 26 zusammenfassend wieder.

Tabelle 26: Gemittelte Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form B über alle 4 Messzeitpunkte

	Gemittelte Itemschwierigkeit p_i				Gemittelte Trennschärfe r_{it}			
	1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
OG	.53	.66	.68	.75	.40	.37	.37	.36
ELG	.84	.90	.91	.93	.30	.28	.23	.24
ELO	.85	.88	.89	.93	.25	.24	.20	.25
ERG	.37	.50	.56	.68	.45	.38	.36	.30

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, ERO = erweitert orthographisches Können

Zusammenfassend betrachtet, enthält die Form A über alle 4 Messzeitpunkte Items mit mittlerer und niedriger Schwierigkeit, die Form B Items mit niedriger bis hoher Schwierigkeit. Die Items wurden von Messzeitpunkt zu Messzeitpunkt leichter und die Schüler schrieben mehr Wörter und Lupenstellen orthographisch richtig. Für die Trennschärfe zeigt sich zusammenfassend, dass die FE-RS 2 Items mit mittlerer und niedriger Trennschärfe enthält.

H_{1.3}: Für die Itemschwierigkeiten des orthographischen Gesamtwerts und der Subskalen der Form A und der Form B gilt zu allen 4 Messzeitpunkten $.2 \leq p_i \leq .8$.

Insgesamt zeigt sich, dass die Schwierigkeitsindexe der FE-RS 2 von Messzeitpunkt zu Messzeitpunkt steigen und die Schüler mehr Wörter und Lupenstellen richtig schreiben. Die gemittelten Itemschwierigkeiten des orthographischen Gesamtwerts und der ERG-Subskala der Form A liegen zu allen 4 Messzeitpunkten im mittleren Bereich zwischen $.2 \leq p_i \leq .8$. Die ELG- und ELO-Subskalen sind über alle 4 Messzeitpunkte sehr leicht; $p_i \geq .8$. Die Hypothese H_{1.3} lässt sich deshalb nur zum Teil bestätigen. Der orthographische Gesamtwert und die ERG-Subskala können Schüler mit unterschiedlichen Rechtschreibleistungen relativ gut differenzieren. Die sehr leichten Items der ELG- und ELO-Subskalen sind dazu geeignet, im unteren Leistungsbereich zu differenzieren. Sie entsprechen damit dem Ziel der FE-RS 2 und ermöglichen differenzierte Aussagen zur alphabetischen Strategie, die die zweite Klassenstufe im Rechtschreiben dominiert.

H_{1.4}: Für die Trennschärfen des orthographischen Gesamtwerts und der Subskalen der Form A und der Form B gilt zu allen 4 Messzeitpunkten $.3 \leq r_{it} \leq .5$.

Die gemittelten Trennschärfen des orthographischen Gesamtwerts und der ERG-Subskala der FE-RS 2 liegen zu allen 4 Messzeitpunkten im mittleren Bereich zwischen $.3 \leq r_{it} \leq .5$. Die gemittelten Trennschärfen der ELG- und ELO-Subskalen sind überwiegend niedrig, $r_{it} \leq .3$. Die Hypothese $H_{1.4}$ lässt sich deshalb nur zum Teil bestätigen. Folglich enthalten allein der orthographische Gesamtwert und die ERG-Subskala trennscharfe Items, die Schüler mit guten und schlechten Rechtschreibleistungen voneinander trennen können. Da aber die Differenzierungsfähigkeit der FE-RS 2 vor allem im unteren Leistungsbereich liegen soll, müssen die Itemschwierigkeiten niedrig sein, was dann zu insgesamt niedrigen Trennschärfen führt (Bortz & Döring, 2006). Insofern sind die Trennschärfen der FE-RS 2 zufriedenstellend.

5.2.2 Reliabilität

Für die Reliabilitätsanalyse der FE-RS 2 wurden die Retest- und Testhalbierungsreliabilitäten, die internen Konsistenzen und die Paralleltestreliabilität berechnet. Die Kennwerte der Paralleltestreliabilität werden vor dem Hintergrund der Forschungsfrage zur Äquivalenz der Form A und Form B im Abschnitt 5.2.5 dargelegt.

5.2.2.1 Retestreliabilität

Für die Retestreliabilitäten der FE-RS 2 wurden die Korrelationskoeffizienten nach Pearson und Spearman zwischen dem 1. und 3. Messzeitpunkt sowie zwischen dem 2. und 4. Messzeitpunkt berechnet. Für die Form A bedeutet das, dass zum 1. und 3. Messzeitpunkt der ersten Gruppe von Schulklassen ($A_1B_2A_3B_4$) und zum 2. und 4. Messzeitpunkt der zweiten Gruppe von Schulklassen ($B_1A_2B_3A_4$) die Form A im Retestintervall von je 20 Wochen vorgelegt wurde. Für die Form B bedeutet das, dass zum 1. und 3. Messzeitpunkt der zweiten Gruppe von Schulklassen ($B_1A_2B_3A_4$) und zum 2. und 4. Messzeitpunkt der ersten Gruppe von Schulklassen ($A_1B_2A_3B_4$) die Form B im Retestintervall von ebenfalls je 20 Wochen vorgelegt wurde. Die Korrelationen für die Retestreliabilität liegen für den orthographischen Gesamtwert im hohen, für die Subskalen überwiegend im mittleren Bereich. Die Retestreliabilität der FE-RS 2 ist damit zufriedenstellend. Tabelle 27 gibt die Korrelationen für die Retestreliabilität der FE-RS 2

für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte wieder.

Tabelle 27: Retestreliabilität der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte

Form A	N	r_{tt}	Form B	N	r_{tt}
OG _{A1→A3}	193	.72 ^{**} (P)	OG _{B1→B3}	232	.71 ^{**} (P)
OG _{A2→A4}	232	.79 ^{**} (S)	OG _{B2→B4}	193	.72 ^{**} (S)
ELG _{A1→A3}	193	.55 ^{**} (S)	ELG _{B1→B3}	232	.64 ^{**} (S)
ELG _{A2→A4}	232	.62 ^{**} (S)	ELG _{B2→B4}	193	.51 ^{**} (S)
ELO _{A1→A3}	193	.48 ^{**} (S)	ELO _{B1→B3}	232	.56 ^{**} (S)
ELO _{A2→A4}	232	.47 ^{**} (S)	ELO _{B2→B4}	193	.49 ^{**} (S)
ERG _{A1→A3}	193	.61 ^{**} (S)	ERG _{B1→B3}	232	.55 ^{**} (S)
ERG _{A2→A4}	232	.67 ^{**} (S)	ERG _{B2→B4}	193	.59 ^{**} (S)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, A₁ = Form A zum 1. Messzeitpunkt, A₂ = Form A zum 2. Messzeitpunkt, A₃ = Form A zum 3. Messzeitpunkt, A₄ = Form A zum 4. Messzeitpunkt, B₁ = Form B zum 1. Messzeitpunkt, B₂ = Form B zum 2. Messzeitpunkt, B₃ = Form B zum 3. Messzeitpunkt, B₄ = Form B zum 4. Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, N = Anzahl, r_{tt} = Retest-Korrelation, ** = $p \leq 0.01$ (2-seitig), (P) = Korrelationskoeffizient nach Pearson, (S) = Korrelationskoeffizient nach Spearman

H_{2.1}: Für die Retestreliabilität des orthographischen Gesamtwerts der Form A und Form B gilt $r_{tt} > .70$. Für die Retestreliabilität der Subskalen der Form A und Form B gilt $.50 < r_{tt} \leq .70$.

Die Retestreliabilität des orthographischen Gesamtwerts beider Formen erreicht mit $r_{tt} = .71$ und $r_{tt} = .79$ die höchsten Werte, die Werte liegen im hohen Bereich. Die Subskalen liegen mit Werten $r_{tt} = .47$ bis $r_{tt} = .67$ im niedrigen bis mittleren Bereich. Die Korrelationen für die ELG- und ERG-Subskalen sind mittelhoch, die Korrelationen für die ELO-Subskalen überwiegend niedrig. Die ELO-Subskala erreicht aber Werte, die annähernd im mittelhohen Bereich liegen. Die Hypothese H_{2.1} lässt sich damit zum Teil bestätigen. Die FE-RS 2 kann überwiegend genau und exakt messen. Vor allem vor dem Hintergrund des sehr langen Retestintervalls von 20 Wochen und der Forderung nach mittelhohen Retestreliabilitäten für veränderungssensitive Verfahren, ist die Retestreliabilität der FE-RS 2 zufriedenstellend.

5.2.2.2 Testhalbierungsreliabilität

Für die Testhalbierungsreliabilität der FE-RS 2 wurden die Items des orthographischen Gesamtwerts und der Subskalen der Form A und Form B aufgrund der komplexer werdenden Silbenstruktur und des steigenden Schwierigkeitsgrades des Wortmaterials nach annähernd gleicher Itemschwierigkeit und Trennschärfe zu Itemzwillingen zusammengestellt. Die Korrelationen zwischen den Testhälften wurden nach der Spearman-Brown-Formel für ungleich und gleich lange Testhälften berechnet. Die Testhalbierungs-Korrelationen liegen für den orthographischen Gesamtwert im hohen und für die Subskalen mit Werten von $r_{tt} = .57$ bis $r_{tt} = .82$ im mittleren bis hohen Bereich. Das heißt, die FE-RS 2 misst zufriedenstellend genau und exakt. Tabelle 28 gibt die Korrelationen für die Testhalbierungsreliabilität der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte wieder.

Tabelle 28: Testhalbierungsreliabilität der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte

	Testhalbierungs-Korrelation r_{tt}							
	OG		ELG		ELO		ERG	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
1. MZP	.87	.89	.68	.73	.72	.74	.64	.80
2. MZP	.87	.87	.78	.74	.79	.82	.76	.73
3. MZP	.86	.85	.71	.57	.74	.74	.60	.67
4. MZP	.84	.86	.65	.71	.64	.65	.74	.61
M	.86	.87	.70	.69	.72	.74	.68	.70

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = 193 für Form A zum 1. und 3. Messzeitpunkt sowie für Form B zum 2. und 4. Messzeitpunkt, N = 232 für Form A zum 2. und 4. Messzeitpunkt sowie für Form B zum 1. und 3. Messzeitpunkt, MZP = Messzeitpunkt, M = Mittelwert, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

H_{2.2}: Für die Testhalbierungsreliabilität des orthographischen Gesamtwerts der Form A und der Form B gilt $r_{tt} > .70$. Für die Testhalbierungsreliabilität der Subskalen der Form A und Form B gilt $.50 < r_{tt} \leq .70$.

Die Werte für die Testhalbierungsreliabilität des orthographischen Gesamtwerts liegen im hohen, die Werte der Subskalen der FE-RS 2 auch überwiegend im hohen Bereich. Die Hypothese H_{2.2} lässt sich deshalb bestätigen. Die FE-RS 2 ist demnach reliabel.

5.2.2.3 Interne Konsistenz

Für die Berechnung der internen Konsistenz der FE-RS 2 wurde der Alphakoeffizient von Cronbach herangezogen. Der Alphakoeffizient wurde für den orthographischen Gesamtwert und die Subskalen berechnet. Insgesamt liegen die internen Konsistenzen für den orthographischen Gesamtwert mit gemittelten Werten von $r_{tt} = .85$ und $r_{tt} = .86$ im hohen Bereich. Die Subskalen erreichen mit gemittelten Werten von $r_{tt} = .65$ bis $r_{tt} = .72$ überwiegend mittelhohe Korrelationen. Tabelle 29 gibt die einzelnen Alphakoeffizienten der Form A und der Form B für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte wieder.

Tabelle 29: Interne Konsistenz der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte

	Interne Konsistenz (Alphakoeffizient)							
	OG		ELG		ELO		ERG	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
1. MZP	.84	.88	.70	.74	.68	.75	.65	.73
2. MZP	.87	.87	.77	.72	.73	.75	.74	.67
3. MZP	.86	.86	.72	.64	.69	.68	.62	.64
4. MZP	.85	.85	.61	.68	.68	.69	.67	.57
M	.85	.86	.70	.69	.70	.72	.67	.65

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = 193 für Form A zum 1. und 3. Messzeitpunkt sowie für Form B zum 2. und 4. Messzeitpunkt, N = 232 für Form A zum 2. und 4. Messzeitpunkt sowie für Form B zum 1. und 3. Messzeitpunkt, MZP = Messzeitpunkt, M = Mittelwert, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

H_{2.3}: Für die interne Konsistenz des orthographischen Gesamtwerts der Form A und der Form B gilt $r_{tt} > .70$. Für die interne Konsistenz der Subskalen der Form A und Form B gilt $.50 < r_{tt} \leq .70$.

Die Werte für den orthographischen Gesamtwert liegen im hohen, die Werte für die Subskalen überwiegend im mittleren Bereich. Die Hypothese H_{2.3} lässt sich deshalb bestätigen. Sowohl der orthographische Gesamtwert als auch die Subskalen der FE-RS 2 messen genau und exakt.

5.2.3 Validität

In diesem Abschnitt werden die Ergebnisse der Validitätsanalyse für die FE-RS 2 dargelegt. Zuerst folgen Aussagen zur Inhaltsvalidität, anschließend zur Konstrukt- und Kriteriumsvalidität.

5.2.3.1 Inhaltsvalidität

H_{3.1}: Die FE-RS 2 ist inhaltsvalide.

Die FE-RS 2 ist ein inhaltlich valides Verfahren zur Erfassung der Rechtschreibleistung im 2. Schuljahr, da es auf der Grundlage der Bildungsstandards im Fach Deutsch für den Primarbereich (KMK, 2005), des Rahmenplans Grundschule Deutsch (MBJS et al., 2004) sowie der zugelassenen Sprachbücher aus dem Schulbuchkatalog 2009/2010 für das Unterrichtsfach Deutsch der Jahrgangsstufe 2 in Grundschulen in Mecklenburg-Vorpommern (MBWK, 2009a) entwickelt wurde. Die Aufgaben der FE-RS 2 sind damit repräsentativ für die vermittelten Unterrichtsinhalte im Lernbereich Rechtschreiben der 2. Klassenstufe. Die Hypothese H_{3.1} lässt sich deshalb bestätigen.

5.2.3.2 Konstruktvalidität

In diesem Abschnitt werden die Ergebnisse zur konkurrenten, diskriminaten und faktoriellen Validität der FE-RS 2 dargelegt.

5.2.3.2.1 Konkurrente Validität

Für die konkurrente Validität der FE-RS 2 wurden die Korrelationen nach Pearson und Spearman zwischen den Testergebnissen des orthographischen Gesamtwerts der FE-RS 2, des DERET 1-2+ (Stock & Schneider, 2008a) und der FE-L 2 (Kuhlmann & Hartke, 2009b) berechnet. Zum 1. Messzeitpunkt wurde einer kleineren Teilstichprobe das Testheft Klasse 1 und zum 4. Messzeitpunkt wurde der gesamten Stichprobe das Testheft Klasse 2

des DERET 1-2+ vorgelegt. Außerdem bearbeitete die gesamte Stichprobe zu allen 4 Messzeitpunkten die FE-L 2 (Kuhlmann & Hartke, 2009b).

Die Übereinstimmungen der FE-RS 2 mit dem DERET 1-2+ liegen im mittleren bis hohen Bereich. Das heißt, die FE-RS 2 erfasst valide die Rechtschreibleistungen. Mit der FE-L 2 korreliert die Form A zum 1. Messzeitpunkt gering, zum 4. Messzeitpunkt mittelmäßig. Die Form B korreliert mit der FE-L 2 zum 1. Messzeitpunkt mittel und zum 4. Messzeitpunkt gering. Zusammenfassend betrachtet, ist die FE-RS 2 ein valides Verfahren zur Erfassung der Rechtschreibleistung. Tabelle 30 zeigt die Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2 und den anderen Validierungsverfahren zum 1. und 4. Messzeitpunkt.

Tabelle 30: Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2, dem DERET 1-2+ und der FE-L 2 zum 1. und 4. Messzeitpunkt

	Form A		Form B	
	1. MZP	4. MZP	1. MZP	4. MZP
DERET 1-2+ (N)	-.71 ^{** (P)} (113)	-.81 ^{** (S)} (232)	-.77 ^{** (P)} (154)	-.69 ^{** (S)} (193)
FE-L 2 (N)	.47 ^{** (P)} (193)	.68 ^{** (S)} (232)	.68 ^{** (P)} (232)	.50 ^{** (S)} (193)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), FE-L 2 = Formative Erfassung der Lesefertigkeit im 2. Schuljahr (Kuhlmann & Hartke, 2009b), N = Anzahl, MZP = Messzeitpunkt, ** = $p \leq 0.01$ (2-seitig), (P) = Korrelationskoeffizient nach Pearson, (S) = Korrelationskoeffizient nach Spearman

H_{3.2}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und dem quantitativen Testergebnis des DERET 1-2+ (Stock & Schneider, 2008a) zum 1. und 4. Messzeitpunkt gilt $r > .70$.

Die Korrelationen zwischen der FE-RS 2 und dem DERET 1-2+ (Stock & Schneider, 2008a) liegen, bis auf die Korrelation zwischen der Form B der FE-RS 2 und dem DERET 1-2+ zum 4. Messzeitpunkt, nicht im mittleren Bereich zwischen $.50 < r \leq .70$, sondern im hohen Bereich. Die Korrelation zwischen der Form B und dem DERET 1-2+ zum 4. Messzeitpunkt erreicht mit $r = -.69$ jedoch einen Wert, der sehr nah am hohen Bereich liegt. Die Hypothese H_{3.2} lässt sich jedoch nur zum Teil bestätigen. Die Folge ist, dass der orthographische Gesamtwert der FE-RS 2 besser als angenommen mit der quantitativen Rechtschreibfehleranalyse des DERET 1-2+ übereinstimmt. Demnach kann die FE-RS 2 valide die Rechtschreibleistungen von Zweitklässlern erfassen.

H_{3.3}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und der Anzahl richtig gelesener Wörter der FE-L 2 zum 1. und 4. Messzeitpunkt gilt $.50 < r \leq .70$.

Die Korrelationen zwischen der FE-RS 2 und der Anzahl richtig gelesener Wörter der FE-L 2 (Kuhlmann & Hartke, 2009b) liegen, bis auf die Korrelation mit der Form A zum 1. und der Form B zum 4. Messzeitpunkt, im mittleren Bereich zwischen $.50 < r \leq .70$. Die Korrelationen mit der Form A zum 1. und die der Form B zum 4. Messzeitpunkt liegen im niedrigen Bereich, liegen mit Werten von $r = .47$ und $r = .50$ aber sehr nah an der Grenze zum mittleren Bereich. Die Hypothese H_{3.3} lässt sich dennoch nur zum Teil bestätigen. Insgesamt bekräftigen die Ergebnisse aber die empirischen Befunde anderer Studien zum Zusammenhang zwischen dem Lesen und Rechtschreiben (z. B. Klicpera et al., 2006; Schabmann und Gasteiger-Klicpera, 2006; P. Marx, 2007). So sind das Lesen und Rechtschreiben zwar keine spiegelbildlichen Prozesse, zwischen ihnen besteht aber ein deutlicher Zusammenhang.

5.2.3.2.2 Diskriminante Validität

Für die diskriminante Validität der FE-RS 2 wurden die Korrelationen nach Pearson und Spearman zwischen den Testergebnissen des orthographischen Gesamtwerts der FE-RS 2, des DEMAT 1+ (Krajewski et al., 2002), des DEMAT 2+ (Krajewski et al., 2004) und der FE-AF 2 (Kuhlmann & Hartke, 2009a) berechnet. Zum 1. Messzeitpunkt wurde einer kleineren Teilstichprobe der DEMAT 1+ und zum 4. Messzeitpunkt wurde der gesamten Stichprobe der DEMAT 2+ vorgelegt. Außerdem bearbeitete die gesamte Stichprobe zu allen 4 Messzeitpunkten die FE-AF 2.

Die Korrelationen der FE-RS 2 mit den genannten Verfahren aus dem Lernbereich Mathematik sind sehr gering. Tabelle 31 zeigt die Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2 und den anderen Validierungsverfahren zum 1. und 4. Messzeitpunkt.

Tabelle 31: Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2, dem DEMAT 1+, dem DEMAT 2+ und der FE-AF 2 zum 1. und 4. Messzeitpunkt

	Form A		Form B	
	1. MZP	4. MZP	1. MZP	4. MZP
FE-AF 2 (N)	-.1 ^(P) (193)	.05 ^(S) (195)	.07 ^(P) (195)	.02 ^(S) (193)
DEMAT 1+ (N)	-.02 ^(P) (124)		-.06 ^(P) (97)	
DEMAT 2+ (N)		-.02 ^(S) (188)		.03 ^(S) (189)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, FE-AF 2 = Formative Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr (Kuhlmann & Hartke, 2009a), DEMAT 1+ = Deutscher Mathematiktest für erste Klassen (Krajewski et al., 2002), DEMAT 2+ = Deutscher Mathematiktest für zweite Klassen (Krajewski et al., 2004), N = Anzahl, MZP = Messzeitpunkt, ** = $p \leq 0.01$ (2-seitig), (P) = Korrelationskoeffizient nach Pearson, (S) = Korrelationskoeffizient nach Spearman

H_{3.4}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und dem quantitativen Testergebnis der FE-AF 2 (Kuhlmann & Hartke, 2009a) zum 1. und 4. Messzeitpunkt gilt $r \leq .50$.

H_{3.5}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und dem Testergebnis des DEMAT 1+ (Krajewski et al., 2002) zum 1. Messzeitpunkt und des DEMAT 2+ (Krajewski et al., 2004) zum 4. Messzeitpunkt gilt $r \leq .50$.

Die Korrelationen zwischen der FE-RS 2, der FE-AF 2 (Kuhlmann & Hartke, 2009a), des DEMAT 1+ (Krajewski et al., 2002) und des DEMAT 2+ (Krajewski et al., 2004) liegen nicht im niedrigen, sondern im sehr niedrigen Bereich. Die Hypothesen H_{3.4} und H_{3.5} lassen sich deshalb bestätigen. Folglich erfasst der FE-RS 2 keine arithmetischen Leistungen. Zudem zeigen die sehr niedrigen Korrelationen, dass Rechtschreiben und mathematische Fähigkeiten in der 2. Klassenstufe unterschiedliche Leistungsbereiche sind, die in keinem Zusammenhang stehen.

5.2.3.2.3 Faktorielle Validität der Form A

Für die Prüfung der Konstruktvalidität der Form A wurden explorative Faktorenanalysen mit den 31 Items zu allen 4 Messzeitpunkten mittels der Hauptkomponentenanalyse durchgeführt. Als Extraktionskriterien wurden das Eigenwertkriterium > 1 , der Scree-Plot

und das hypothetische Modell gewählt. Für das hypothetische Modell wurden entsprechend des der FE-RS 2 zugrunde gelegten Kategoriensystems vier zu extrahierende Faktoren vorgegeben. Als Rotationstechniken wurden für eine erste Analyse die Varimax-Rotation und für eine zweite Analyse die Promax-Rotation gewählt, um zu entscheiden, ob die Faktorenstruktur methodeninvariant ist (Bühner, 2004). Vor der Durchführung der Faktorenanalysen wurden deren Voraussetzungen für die Berechnung geprüft. Für alle 4 Messzeitpunkte lehnen die Bartlett-Tests mit einem Signifikanzniveau von $p \leq 0.001$ die Hypothese ab, dass alle Itemkorrelationen in der Population gleich Null sind. Die KMO-Koeffizienten bestätigen für den 1. Messzeitpunkt mit einem Wert von $KMO = .77$ eine mittlere Eignung, für den 2. Messzeitpunkt mit $KMO = .86$, für den 3. Messzeitpunkt mit $KMO = .80$ und für den 4. Messzeitpunkt mit $KMO = .82$ eine gute Eignung (Bühner, 2004) der Items für die Faktorenanalysen. Auch die MSA-Koeffizienten, die sich für alle 4 Messzeitpunkte überwiegend im mittleren bis guten Bereich befinden, rechtfertigen die Durchführung der Faktorenanalyse. Die einzelnen MSA-Koeffizienten für die Items der Form A über alle 4 Messzeitpunkte befinden sich in Tabelle N1 im Anhang N. Die Voraussetzungen für die Berechnung der Faktorenanalysen der Form A über alle 4 Messzeitpunkte sind erfüllt.

Zum 1. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 10 Faktoren, die zusammen 59 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 6 Faktoren, die zusammen 45 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 18 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle O1 und der Scree-Plot in Abbildung O1 im Anhang O dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen O2 und O3 im Anhang O zeigen die rotierte orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) 10-Faktorenstruktur für die FE-RS 2 A zum 1. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Die Lösung mit 4 Faktoren erklärt 36 % der Gesamtvarianz. Beide Rotationstechniken führen zu einer gleichen Zuordnung der Items zu den Faktoren. Die Tabellen O4 und O5 im Anhang O zeigen die rotierte orthogonale

und oblique 4-Faktorenstruktur der Form A zum 1. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zum 2. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 9 Faktoren, die zusammen 58 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 5 Faktoren, die zusammen 44 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 23 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle P1 und der Scree-Plot in Abbildung P1 im Anhang P dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen P2 und P3 im Anhang P zeigen die orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) 9-Faktorenstruktur für die Form A zum 2. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden erneut Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Die Lösung mit 4 Faktoren erklärt 39 % der Gesamtvarianz. Beide Rotationstechniken führen zu einer gleichen Zuordnung der Items zu den Faktoren. Die Tabellen P4 und P5 im Anhang P zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form A zum 2. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zum 3. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 10 Faktoren, die zusammen 59 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 2 oder 4 Faktoren, die zusammen 26 % bzw. 36 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 20 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle Q1 und der Scree-Plot in Abbildung Q1 im Anhang Q dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen Q2 und Q3 im Anhang Q zeigen die orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) 10-Faktorenstruktur für die Form A zum 3. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden erneut Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Die

Lösung mit 4 Faktoren erklärt 36 % der Gesamtvarianz. Beide Rotationstechniken führen zu einer annähernd gleichen Zuordnung der Items zu den Faktoren. Die Tabellen Q4 und Q5 im Anhang Q zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form A zum 3. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zum 4. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 10 Faktoren, die zusammen 58 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 3 Faktoren, die zusammen 31 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 19 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle R1 und der Scree-Plot in Abbildung R1 im Anhang R dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen R2 und R3 im Anhang R zeigen die orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) Faktorenstruktur für die Form A zum 4. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden erneut Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Die Lösung mit 4 Faktoren erklärt 35 % der Gesamtvarianz. Beide Rotationstechniken führen zu einer gleichen Zuordnung der Items zu den Faktoren. Die Tabellen R4 und R5 im Anhang R zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form A zum 4. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zusammenfassend betrachtet, zeigen sich für die Form A nach dem Eigenwertkriterium für jeden Messzeitpunkt unabhängig von der Rotationstechnik annähernd stabile Zuordnungen der Items zu den Faktoren. Die Zuordnungen der Items zu den Faktoren sind aber nicht über alle 4 Messzeitpunkte stabil. Lediglich die Items *reden*, *fragen*, *schlafen*, *flach*, *hart* und *singen* laden über alle 4 Messzeitpunkte stabil auf dem ersten Faktor. Nach dem hypothetischen Modell zeigen sich für die Form A unabhängig von der Rotationstechnik für jeden Messzeitpunkt, bis auf den 3. Messzeitpunkt, stabile Zuordnungen der Items zu den Faktoren. Die Zuordnungen der Items zu den Faktoren sind aber erneut nicht über alle 4 Messzeitpunkte stabil. Lediglich die Items *Nase*, *reden*, *Krone*, *fragen*, *schlafen*, *flach*, *hart*, *helfen*, *Leute* und *singen* laden über alle 4 Messzeitpunkte auf dem gleichen Faktor. Dennoch zeigen sich relativ eindeutige Zuordnungen der Items zu den Faktoren über alle 4

Messzeitpunkte. Auf dem Faktor 1 laden überwiegend Verben und Adjektive, die durch Konsonantenhäufung und die Endsilbe <-en> charakterisiert sind. Der Faktor 1 enthält folglich Wörter, die der ELG- und ELO-Subskala zugeordnet werden. Auf dem Faktor 2 laden überwiegend Wörter mit Kürze- und Dehnungszeichen, weshalb der Faktor 2 der ERG-Subskala zugeordnet werden kann. Auf dem Faktor 3 laden überwiegend Nomen, die durch Konsonantenhäufung charakterisiert sind. Der Faktor 3 enthält folglich Wörter, die der ELG- und ELO-Subskala zugeordnet werden. Auf dem Faktor 4 laden überwiegend alphabetische Schreibungen von Wörtern, was für eine Zuordnung zur ELG-Subskala spricht. Zusammenfassend betrachtet, können die Faktorenanalysen nach dem hypothetischen Modell die vier Faktoren des zugrunde gelegten Kategoriensystems der FE-RS 2 nicht bestätigen. Die Analysen lassen jedoch eine eindeutige und stabile Zuordnung der Items zu zwei Faktoren zu, die als elementare und erweiterte Stufe bezeichnet werden. Das heißt, die ermittelte Struktur bestätigt die Grundstruktur der Form A und damit dem dieser Arbeit zugrunde liegenden Entwicklungsmodell des graphematischen und orthographischen Rechtschreibkönnens.

5.2.3.2.4 Faktorielle Validität der Form B

Für die Prüfung der Konstruktvalidität der Form B wurden explorative Faktorenanalysen mit den 31 Items zu allen 4 Messzeitpunkten mittels der Hauptkomponentenanalyse durchgeführt. Als Extraktionskriterien wurden das Eigenwertkriterium > 1 , der Scree-Plot und das hypothetische Modell gewählt. Für das hypothetische Modell wurden entsprechend des der FE-RS 2 zugrunde gelegten Kategoriensystems vier zu extrahierenden Faktoren vorgegeben. Als Rotationstechniken wurden für eine erste Analyse die Varimax-Rotation und für eine zweite Analyse die Promax-Rotation gewählt, um zu entscheiden, ob die Faktorenstruktur methodeninvariant ist (Bühner, 2004). Vor der Durchführung der Faktorenanalysen wurden deren Voraussetzungen für die Berechnung geprüft. Für alle 4 Messzeitpunkte lehnen die Bartlett-Tests mit einem Signifikanzniveau von $p \leq 0.001$ die Hypothese ab, dass alle Itemkorrelationen in der Population gleich Null sind. Die KMO-Koeffizienten bestätigen für den 1. Messzeitpunkt mit einem Wert von $KMO = .86$, für den 2. und 3. Messzeitpunkt mit $KMO = .82$ und für den 4. Messzeitpunkt mit $KMO = .80$ eine gute Eignung (Bühner, 2004) der Items für die Faktorenanalysen. Auch die MSA-Koeffizienten, die sich für alle 4 Messzeitpunkte überwiegend im mittleren

bis guten Bereich befinden, rechtfertigen die Durchführung der Faktorenanalyse. Die einzelnen MSA-Koeffizienten für die Items der Form B über alle 4 Messzeitpunkte befinden sich in Tabelle S1 im Anhang S. Die Voraussetzungen für die Berechnung der Faktorenanalysen der Form B über alle 4 Messzeitpunkte sind erfüllt.

Zum 1. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 8 Faktoren, die zusammen 55 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 6 Faktoren, die zusammen 48 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 23 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle T1 und der Scree-Plot in Abbildung T1 im Anhang T dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu gleichen Zuordnungen der Items zu Faktoren und damit zu einer stabilen Faktorenstruktur. Die Tabellen T2 und T3 im Anhang T zeigen die rotierte orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) 10-Faktorenstruktur für die Form B zum 1. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Die Lösung mit 4 Faktoren erklärt 40 % der Gesamtvarianz. Beide Rotationstechniken führen zu einer annähernd gleichen Zuordnung der Items zu den Faktoren. Die Tabellen T4 und T5 im Anhang T zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form B zum 1. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zum 2. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 10 Faktoren, die zusammen 59 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 4 Faktoren, die zusammen 37 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 21 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle U1 und der Scree-Plot in Abbildung U1 im Anhang U dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen U2 und U3 im Anhang U zeigen die orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) 10-Faktorenstruktur für die Form B zum 2. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-

Faktorenlösung widersprechen, wurden erneut Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Beide Rotationstechniken führen zu einer annähernd gleichen Zuordnung der Items zu den Faktoren. Die Tabellen U4 und U5 im Anhang U zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form B zum 2. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zum 3. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 9 Faktoren, die zusammen 56 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 2 oder 4 Faktoren, die zusammen 28 % bzw. 37 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 21 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle V1 und der Scree-Plot in Abbildung V1 im Anhang V dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen V2 und V3 im Anhang V zeigen die orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) 9-Faktorenstruktur für die Form B zum 3. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden erneut Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Beide Rotationstechniken führen zu einer gleichen Zuordnung der Items zu den Faktoren. Die Tabellen V4 und V5 im Anhang V zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form B zum 3. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zum 4. Messzeitpunkt zeigt sich nach dem Eigenwertkriterium eine Lösung mit 11 Faktoren, die zusammen 63 % der Gesamtvarianz erklären, und nach dem Scree-Plot eine Lösung mit 6 Faktoren, die zusammen 45 % der Gesamtvarianz erklären. Der erste Faktor klärt bereits 20 % der Gesamtvarianz auf, alle weiteren Faktoren weniger. Die Eigenwerte der Faktoren und die erklärte Gesamtvarianz sind in Tabelle W1 und der Scree-Plot in Abbildung W1 im Anhang W dargestellt. Die Varimax-Rotation und Promax-Rotation führen zu annähernd gleichen Zuordnungen der Items zu Faktoren und damit zu einer annähernd stabilen Faktorenstruktur. Die Tabellen W2 und W3 im Anhang W zeigen die orthogonale (Varimax-Rotation) und oblique (Promax-Rotation) Faktorenstruktur für die

Form B zum 4. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind. Da diese Faktorenlösungen und Faktorenstrukturen der erwarteten theoriebasierten 4-Faktorenlösung widersprechen, wurden erneut Hauptkomponentenanalysen nach dem hypothetischen Modell mit Varimax-Rotation und Promax-Rotation durchgeführt. Die Lösung mit 4 Faktoren erklärt 36 % der Gesamtvarianz. Beide Rotationstechniken führen zu einer annähernd gleichen Zuordnung der Items zu den Faktoren. Die Tabellen W4 und W5 im Anhang W zeigen die rotierte orthogonale und oblique 4-Faktorenstruktur der Form B zum 4. Messzeitpunkt, wobei nur Faktorladungen $> .30$ aufgeführt sind.

Zusammenfassend betrachtet, zeigen sich für die Form B nach dem Eigenwertkriterium für jeden Messzeitpunkt unabhängig von der Rotationstechnik annähernd stabile Zuordnungen der Items zu den Faktoren. Die Zuordnungen der Items zu den Faktoren sind aber nicht über alle 4 Messzeitpunkte stabil. Lediglich die Items *legen, tragen, schlagen, frech, kalt, halten, Quark, fangen* und *Tatze* laden über alle 4 Messzeitpunkte stabil auf einem Faktor. Nach dem hypothetischen Modell zeigen sich für die Form B unabhängig von der Rotationstechnik für jeden Messzeitpunkt annähernd stabile Zuordnungen der Items zu den Faktoren. Die Zuordnungen der Items zu den Faktoren sind aber erneut nicht über alle 4 Messzeitpunkte stabil. Lediglich die Items *Opa, Seife, legen, tragen, schlagen, frech, kalt, halten* und *fangen* laden über alle 4 Messzeitpunkte auf dem gleichen Faktor. Die Deutung der Faktoren gelingt nicht so eindeutig wie bei der Form A. Allein der Faktor 1 kann dahingehend interpretiert werden, dass auf diesem Faktor wie bei der Form A überwiegend Verben und Adjektive laden, die durch Konsonantenhäufung und die Endsilbe ⟨-en⟩ charakterisiert sind. Der Faktor 1 enthält folglich Wörter, die der ELG- und ELO-Subskala zugeordnet werden. Alle anderen Faktoren sind für eine Deutung zu instabil.

H_{3,6}: Die Form A und Form B erfassen die Subskalen ELG, ERG, ELO und ERO. Es zeigen sich stabile Zuordnungen der Items der Subskalen zu Faktoren über alle 4 Messzeitpunkte.

Die FE-RS 2 erfasst nicht die vier Subskalen ELG, ERG, ELO und ERO. Die Hypothese H_{3,6} lässt sich deshalb nicht bestätigen. Die Analysen für die Form A lassen aber die Grundstruktur der FE-RS 2, eine elementare und erweiterte Stufe, erkennen und entsprechen damit dem Rechtschreibentwicklungsmodell der vorliegenden Arbeit. Für die Form B ist aufgrund der Analysen diese Grundstruktur, bis auf die elementare Stufe, nicht

erkennbar. Die Ergebnisse stützen damit nur zum Teil die Konstruktvalidität der FE-RS 2.

5.2.3.3 Übereinstimmungsvalidität

Für die Übereinstimmungsvalidität der FE-RS 2 wurden die Korrelationen nach Spearman zwischen den Testergebnissen des orthographischen Gesamtwerts der FE-RS 2 und den Rechtschreibnoten der Schüler zum 4. Messzeitpunkt berechnet. Es ist zu beachten, dass eine Schule die Noten ihrer 61 Zweitklässler nicht bekannt gab und 12 Schüler aufgrund von Notenbefreiungen aus der Gesamtstichprobe fielen.

Die Übereinstimmungen der FE-RS 2 mit der Rechtschreibnote liegen für die Form B im mittleren und für die Form A im hohen Bereich. Das bedeutet, dass die FE-RS 2 eng am Rahmenplan und an den vermittelten Unterrichtsinhalten des Lernbereichs Rechtschreiben ausgerichtet ist. Tabelle 32 zeigt die Korrelation zwischen dem orthographischen Gesamtwert der FE-RS 2 und der Rechtschreibnote zum 4. Messzeitpunkt.

Tabelle 32: Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2 und der Rechtschreibnote zum 4. Messzeitpunkt

	Form A		Form B	
	1. MZP	4. MZP	1. MZP	4. MZP
Rechtschreibnote (N)		-.72 ^{** (S)} (175)		-.62 ^{** (S)} (177)
Anmerkungen:	FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, MZP = Messzeitpunkt, ** = $p \leq 0.01$ (2-seitig), (S) = Korrelationskoeffizient nach Spearman			

H_{3.7}: Für die Korrelationen zwischen dem orthographischen Gesamtwert der Form A sowie der Form B und den Rechtschreibnoten der Schüler zum 4. Messzeitpunkt gilt $r > .70$.

Die Korrelationen zwischen der FE-RS 2 und den Rechtschreibnoten der Schüler liegen zum 4. Messzeitpunkt für die Form A im hohen und für die Form B im mittleren Bereich. Die Hypothese H_{3.7} lässt sich deshalb nur zum Teil bestätigen. Die Korrelationen bestätigen, dass die FE-RS 2 valide die Rechtschreibleistungen von Zweitklässlern erfassen kann.

5.2.3.4 Ergebnisse der Regressionsanalyse zur Schätzung der prognostischen Validität

Um Aussagen zur prognostischen Validität der FE-RS 2 treffen zu können, wurden für jede Form der FE-RS 2 zwei lineare Regressionsanalysen durchgeführt. Für beide Formen der FE-RS 2 diente die erste Regressionsanalyse dazu, den orthographischen Gesamtwert zum 3. Messzeitpunkt (Kriterium) durch den orthographischen Gesamtwert des 1. Messzeitpunktes (Prädiktor) vorherzusagen. Die zweite Regressionsanalyse diente für beide Formen der FE-RS 2 dazu, den orthographischen Gesamtwert zum 4. Messzeitpunkt durch den orthographischen Gesamtwert des 2. Messzeitpunktes vorherzusagen. Vor der Durchführung der Regressionsanalysen wurden deren Voraussetzungen für die Berechnung geprüft. Wie in Abbildung X1 im Anhang X zu erkennen, bestehen zwischen den Prädiktoren sowie zwischen den Prädiktoren und Kriterien lineare Zusammenhänge. Es gibt keine Ausreißerwerte. Somit kann von Linearität ausgegangen werden. Die Prädiktoren korrelieren mit den Kriterien mittel bzw. hoch. Tabelle 33 gibt die Korrelationen zwischen den orthographischen Gesamtwerten der FE-RS 2 wieder.

Tabelle 33: Korrelationen zwischen den orthographischen Gesamtwerten der FE-RS 2

	OG _{A3}	OG _{B4}	OG _{B3}	OG _{A4}
OG _{A1} (N)	.71 ^{** (S)} (193)			
OG _{B2} (N)		.72 ^{** (S)} (193)		
OG _{B1} (N)			.69 ^{** (S)} (232)	
OG _{A2} (N)				.79 ^{** (P)} (232)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, OG = orthographischer Gesamtwert, A₁ = Form A zum 1. Messzeitpunkt, A₂ = Form A zum 2. Messzeitpunkt, A₃ = Form A zum 3. Messzeitpunkt, A₄ = Form A zum 4. Messzeitpunkt, B₁ = Form B zum 1. Messzeitpunkt, B₂ = Form B zum 2. Messzeitpunkt, B₃ = Form B zum 3. Messzeitpunkt, B₄ = Form B zum 4. Messzeitpunkt, N = Anzahl, ** = $p \leq 0.01$ (2-seitig), (P) = Korrelationskoeffizient nach Pearson, (S) = Korrelationskoeffizient nach Spearman

Streng genommen, kann für alle drei Regressionsanalysen nicht von Homoskedastizität ausgegangen werden, da, wie in Abbildung X2 im Anhang X zu erkennen, die geschätzten Werte nicht gleichmäßig um die Nulllinien streuen. Ebenso kann nicht für alle Residuen der vier Regressionsanalysen von einer augenscheinlichen Normalverteilung ausgegangen werden. „Das bedeutet, dass die Angaben aus den Signifikanztests (Regressionsgewichte und R) nicht korrekt sind“ (Bühner & Ziegler, 2009, S. 673). Die Häufigkeitsverteilungen der Residuen der FE-RS 2 sind in Abbildung X3 im Anhang X dargestellt. Bis auf das

Kriterium der Homoskedastizität und zum Teil das Kriterium der Normalverteilung der Residuen sind damit die Voraussetzungen für die Berechnung der Regressionsanalyse erfüllt.

Nachfolgend werden die zwei Regressionsanalysen für die FE-RS 2 der Form A dargestellt. In die erste Regressionsanalyse wurden der orthographische Gesamtwert zum 1. Messzeitpunkt als Prädiktor und der orthographische Gesamtwert zum 3. Messzeitpunkt als Kriterium aufgenommen. Tabelle 34 stellt das Ergebnis der Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form A zum 3. Messzeitpunkt dar.

Tabelle 34: Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form A zum 3. Messzeitpunkt durch den orthographischen Gesamtwert der Form A zum 1. Messzeitpunkt (N = 193)

Variable	R^2_{kor}	B	β	d	CI
Modell 1	.52			1.72	6.94
Orthographischer Gesamtwert 1. MZP		.75	.73***		
Anmerkungen:	Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, R^2_{kor} = korrigiertes R^2 = Bestimmtheitsmaß, B = unstandardisiertes Regressionsgewicht, β = standardisiertes Regressionsgewicht, d = Durbin-Watson-Koeffizient, CI = Konditionsindex, MZP = Messzeitpunkt, *** = $p \leq 0.001$				

Der 1. Messzeitpunkt der Form A klärt 52 % der Varianz des orthographischen Gesamtwerts zum 3. Messzeitpunkt auf. Das standardisierte Regressionsgewicht zeigt, dass der 1. Messzeitpunkt den orthographischen Gesamtwert zum 3. Messzeitpunkt höchst signifikant vorhersagen kann. Der Durbin-Watson-Koeffizient ist als akzeptabel einzuschätzen (Bühl, 2008) und deutet auf keine Autokorrelationen hin. Der Konditionsindex deutet mit $CI < 15$ auf keine Kollinearität hin (Bühner & Ziegler, 2009).

In die zweite Regressionsanalyse wurden der orthographische Gesamtwert zum 2. Messzeitpunkt als Prädiktor und der orthographische Gesamtwert zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabelle 35 stellt das Ergebnis der Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form A zum 4. Messzeitpunkt dar.

Tabelle 35: Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form A zum 4. Messzeitpunkt durch den orthographischen Gesamtwert der Form A zum 2. Messzeitpunkt (N = 232)

Variable	R^2_{korrr}	B	β	d	CI
Modell 1	.62			1.60	6.12
Orthographischer Gesamtwert 2. MZP		.68	.79***		

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, N = Anzahl, R^2_{korrr} = korrigiertes R^2 = Bestimmtheitsmaß, B = unstandardisiertes Regressionsgewicht, β = standardisiertes Regressionsgewicht, d = Durbin-Watson-Koeffizient, CI = Konditionsindex, MZP = Messzeitpunkt, *** = $p \leq 0.001$

Der 2. Messzeitpunkt der Form A klärt 62 % der Varianz des orthographischen Gesamtwerts zum 4. Messzeitpunkt auf. Das standardisierte Regressionsgewicht zeigt, dass der 2. Messzeitpunkt den orthographischen Gesamtwert zum 4. Messzeitpunkt höchst signifikant vorhersagen kann. Der Durbin-Watson-Koeffizient ist als akzeptabel einzuschätzen (Bühl, 2008) und deutet auf keine Autokorrelationen hin. Der Konditionsindex deutet mit $CI < 15$ auf keine Kollinearität hin (Bühner & Ziegler, 2009).

Nachfolgend werden die zwei Regressionsanalysen für die FE-RS 2 der Form B dargestellt. In die erste Regressionsanalyse wurden der orthographische Gesamtwert zum 1. Messzeitpunkt als Prädiktor und der orthographische Gesamtwert zum 3. Messzeitpunkt als Kriterium aufgenommen. Tabelle 36 stellt das Ergebnis der Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form B zum 3. Messzeitpunkt dar.

Tabelle 36: Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form B zum 3. Messzeitpunkt durch den orthographischen Gesamtwert der Form B zum 1. Messzeitpunkt (N = 193)

Variable	R^2_{korrr}	B	β	d	CI
Modell 1	.51			1.83	5.51
Orthographischer Gesamtwert 1. MZP		.66	.72***		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, R^2_{korrr} = korrigiertes R^2 = Bestimmtheitsmaß, B = unstandardisiertes Regressionsgewicht, β = standardisiertes Regressionsgewicht, d = Durbin-Watson-Koeffizient, CI = Konditionsindex, MZP = Messzeitpunkt, *** = $p \leq 0.001$

Der 1. Messzeitpunkt der Form B klärt 51 % der Varianz des orthographischen Gesamtwerts zum 3. Messzeitpunkt auf. Das standardisierte Regressionsgewicht zeigt, dass der 1. Messzeitpunkt den orthographischen Gesamtwert zum 3. Messzeitpunkt höchst signifikant vorhersagen kann. Der Durbin-Watson-Koeffizient ist als akzeptabel einzuschätzen (Bühl, 2008) und deutet auf keine Autokorrelationen hin. Der Konditionsindex deutet mit $CI < 15$ auf keine Kollinearität hin (Bühner & Ziegler, 2009).

In die zweite Regressionsanalyse wurden der orthographische Gesamtwert zum 2. Messzeitpunkt als Prädiktor und der orthographische Gesamtwert zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabelle 37 stellt das Ergebnis der Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form B zum 4. Messzeitpunkt dar.

Tabelle 37: Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form B zum 4. Messzeitpunkt durch den orthographischen Gesamtwert der Form B zum 2. Messzeitpunkt (N = 232)

Variable	$R^2_{\text{kor}} \quad$	B	β	d	CI
Modell 1	.53			1.85	7.27
Orthographischer Gesamtwert 2. MZP		.65	.73***		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, N = Anzahl, R^2_{kor} = korrigiertes R^2 = Bestimmtheitsmaß, B = unstandardisiertes Regressionsgewicht, β = standardisiertes Regressionsgewicht, d = Durbin-Watson-Koeffizient, CI = Konditionsindex, MZP = Messzeitpunkt, *** = $p \leq 0.001$

Der 2. Messzeitpunkt der Form B klärt 53 % der Varianz des orthographischen Gesamtwerts zum 4. Messzeitpunkt auf. Das standardisierte Regressionsgewicht zeigt, dass der 2. Messzeitpunkt den orthographischen Gesamtwert zum 4. Messzeitpunkt höchst signifikant vorhersagen kann. Der Durbin-Watson-Koeffizient ist als akzeptabel einzuschätzen (Bühl, 2008) und deutet auf keine Autokorrelationen hin. Der Konditionsindex deutet mit $CI < 15$ auf keine Kollinearität hin (Bühner & Ziegler, 2009).

Zusammenfassend betrachtet, ist die Varianzaufklärung der Rechtschreibleistung zum 3. Messzeitpunkt durch den 1. Messzeitpunkt für beide Formen der FE-RS 2 mit 51 % und 52 % annähernd gleich hoch. Für die Vorhersage der Rechtschreibleistung zum 4. Messzeitpunkt durch den orthographischen Gesamtwert zum 2. Messzeitpunkt leistet die Form A mit 62 % Varianzaufklärung einen höheren Beitrag als die Form B. Insgesamt ist damit eine relativ frühe und sichere Vorhersage der Rechtschreibleistung am Ende des ersten Schulhalbjahres der 2. Klassenstufe durch den orthographischen Gesamtwert der FE-RS 2 der Form A möglich.

5.2.3.5 Ergebnisse der klassifikatorischen Analyse zur Schätzung der prognostischen Validität

Da die FE-RS 2 als Screeningverfahren Rechtschreibschwierigkeiten von Schülern voraussagen können soll, wurde mittels klassifikatorischer Analysen deren prognostische

Validität überprüft. Die Testkennwerte zur prognostischen Validität (Sensitivität, Spezifität, positiver und negativer prädiktiver Wert, RAZ-Index) wurden mit einer Excel-Auswertungshilfe (Lenhard, 2011) berechnet. Um die Kennwerte zur Vorhersagegüte der FE-RS 2 berechnen zu können, wurden vorab Prädiktoren und Kriterien definiert. Prädiktoren sind aufgrund des zugrunde liegenden Untersuchungsdesigns der orthographische Gesamtwert sowie die Subskalen zum 1. und 2. Messzeitpunkt. Für die Dichotomisierung der Prädiktoren in „Risiko“ und „Kein Risiko“ wurde der Prozentrang $PR \leq 25$ gewählt. Dies wird dem Anspruch eines Screenings gerecht, alle Schüler zu erfassen, die ein Risiko haben, Rechtschreibschwierigkeiten zu entwickeln. Kriterien sind aufgrund des zugrunde liegenden Untersuchungsdesigns der orthographische Gesamtwert zum 3. und 4. Messzeitpunkt, die Rechtschreibnote zum 4. Messzeitpunkt und die Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt. Für die Dichotomisierung der Kriterien in „betroffen“ und „nicht betroffen“ wurde für den orthographischen Gesamtwert der Prozentrang $PR \leq 15$, für die Rechtschreibnote schlechter gleich vier und für den DERET 1-2+ (Stock & Schneider, 2008a) der Prozentrang $PR \leq 16$ gewählt, um die Schüler zu identifizieren, deren Rechtschreibleistungen sich im unterdurchschnittlichen Bereich befinden.

Die Anzahl der Schüler, die aufgrund des orthographischen Gesamtwerts und der Anzahl richtig geschriebener Lupenstellen der Subskalen der FE-RS 2 zu den ersten zwei Messzeitpunkten ein Risiko hat, Rechtschreibschwierigkeiten zu entwickeln, ist in Tabelle 38 aufgeführt.

Tabelle 38: Anzahl der Risikoschüler der FE-RS 2 zum 1. und 2. Messzeitpunkt aufgrund kritischer Prädiktoren-Testwerte ($PR \leq 25$, $N_{A1/B2} = 193$, $N_{B1/A2} = 232$)

Skalen	1. MZP A	1. MZP B	2. MZP A	2. MZP B
OG N (%)	60 (14.1)	71 (16.7)	68 (16.0)	52 (12.2)
ELG N (%)	54 (12.7)	69 (16.2)	63 (14.8)	61 (14.4)
ELO N (%)	63 (14.8)	74 (17.4)	64 (15.1)	62 (14.6)
ERG N (%)	52 (12.2)	75 (17.6)	81 (19.1)	61 (14.4)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, PR = Prozentrang, N = Anzahl, A₁ = Form A zum 1. Messzeitpunkt, B₂ = Form B zum 2. Messzeitpunkt, B₁ = Form B zum 1. Messzeitpunkt, A₂ = Form A zum 2. Messzeitpunkt, MZP = Messzeitpunkt, A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Die Anzahl der Schüler, die aufgrund des orthographischen Gesamtwerts zum 3. und 4. Messzeitpunkt, aufgrund der Rechtschreibnote und der Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt unterdurchschnittliche Rechtschreibleistungen hat und damit zu den Problemschülern zählt, ist in Tabelle 39 aufgeführt.

Tabelle 39: Anzahl der Problemschüler der FE-RS 2 zum 3. und 4. Messzeitpunkt aufgrund kritischer Kriterien-Testwerte ($PR \leq 15$ für OG, $RS\text{-}Note \geq 4$, $PR \leq 16$ für Anzahl der Rechtschreibfehler im DERET 1-2+, $N_{A3/B4} = 193$, $N_{B3/A4} = 232$)

	3. MZP A	3. MZP B	4. MZP A	4. MZP B	4. MZP
OG N (%)	33 (7.8)	43 (10.1)	35 (8.2)	37 (8.7)	
RS-Note N (%)					60 (17.0)*
DERET 1-2+ N (%)					63 (14.8)**
Anmerkungen:	FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, PR = Prozentrang, OG = orthographischer Gesamtwert, RS-Note = Rechtschreibnote, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), N = Anzahl, A ₃ = Form A zum 3. Messzeitpunkt, B ₄ = Form B zum 4. Messzeitpunkt, B ₃ = Form B zum 3. Messzeitpunkt, A ₄ = Form A zum 4. Messzeitpunkt, MZP = Messzeitpunkt, A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, * = Da die Rechtschreibnoten nicht von der Gesamtstichprobe vorliegen, beträgt die Gesamtanzahl diesbezüglich N = 352, ** = N = 425				

5.2.3.5.1 Prognostische Validität der Form A

In die erste klassifikatorische Analyse wurden der orthographische Gesamtwert und die Subskalen zum 1. Messzeitpunkt als Prädiktoren und der orthographische Gesamtwert zum 3. Messzeitpunkt als Kriterium aufgenommen. In die zweite klassifikatorische Analyse wurden der orthographische Gesamtwert und die Subskalen zum 2. Messzeitpunkt als Prädiktoren und der orthographische Gesamtwert zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabellen 40 und 41 stellen die Ergebnisse dieser klassifikatorischen Analysen dar.

Tabelle 40: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 1. Messzeitpunkt und des Kriteriums zum 3. Messzeitpunkt

Prädiktoren zum 1. MZP (PR ≤ 25)	Kriterium: Orthographischer Gesamtwert zum 3. MZP (PR ≤ 15)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	79	79	43	95	69	0.6	3.8	0.6
ELG	67	80	41	92	54	0.5	3.4	0.4
ELO	55	72	29	89	33	0.3	2.0	0.6
ERG	55	79	35	89	38	0.3	2.6	0.6

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 41: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 2. Messzeitpunkt und des Kriteriums zum 4. Messzeitpunkt

Prädiktoren zum 2. MZP (PR ≤ 25)	Kriterium: Orthographischer Gesamtwert zum 4. MZP (PR ≤ 15)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	80	80	41	96	72	0.6	4.0	0.3
ELG	66	80	37	93	53	0.5	3.3	0.4
ELO	60	78	33	92	45	0.4	2.7	0.5
ERG	66	71	28	92	47	0.4	2.3	0.5

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Für den 1. und 2. Messzeitpunkt und das Kriterium des orthographischen Gesamtwerts zum 4. Messzeitpunkt der Form A erreicht die Sensitivität Werte von 55 % bis 80 %, wobei die ELO-Skala am schwächsten und der orthographische Gesamtwert Problemschüler am stärksten identifizieren. Die Spezifität erreicht Werte von 71 % bis 80 %, wobei die ELO- und ERG-Skalen am schwächsten und der orthographische Gesamtwert und die ELG-Skala am stärksten Nicht-Problemschüler identifizieren. Der positive prädiktive Wert liegt zwischen 28 % und 43 %, der negative prädiktive Wert zwischen 89 % und 96 %. Die RAZ-IndeXe liegen mit Werten zwischen 38 % und 72 % im guten bis sehr guten Bereich, wobei die RAZ-IndeXe für den orthographischen

Gesamtwert am höchsten sind. Die Youden-Indexe erreichen Werte von höchstens $Y-I = 0.6$. Die positiven und negativen Likelihood-Ratios deuten auf überwiegend schwache oder kaum relevante diagnostische Evidenz der Form A hin.

In die dritte bis fünfte klassifikatorische Analyse für die Form A wurden der orthographische Gesamtwert und die Subskalen zum 1., 2. und 3. Messzeitpunkt als Prädiktoren und die Rechtschreibnote zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabellen 42, 43 und 44 stellen die Ergebnisse der klassifikatorischen Analysen zur Vorhersage der Rechtschreibnote zum 4. Messzeitpunkt dar.

Tabelle 42: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt

Prädiktoren zum 1. MZP (PR \leq 25)	Kriterium: Rechtschreibnote zum 4. MZP (RS-Note \geq 4)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	63	75	24	94	48	0.4	2.5	0.5
ELG	53	78	22	9	37	0.3	2.4	0.6
ELO	58	72	20	93	39	0.3	2.6	0.6
ERG	63	78	26	95	50	0.4	2.9	0.5

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, RS-Note = Rechtschreibnote, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 43: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt

Prädiktoren zum 2. MZP (PR \leq 25)	Kriterium: Rechtschreibnote zum 4. MZP (RS-Note \geq 4)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	63	87	59	89	51	0.5	4.8	0.4
ELG	59	89	62	88	50	0.5	5.4	0.5
ELO	44	81	41	82	25	0.3	2.3	0.7
ERG	71	75	46	89	54	0.5	2.8	0.4

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, RS-Note = Rechtschreibnote, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 44: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt

Prädiktoren zum 3. MZP (PR \leq 25)	Kriterium: Rechtschreibnote zum 4. MZP (RS-Note \geq 4)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	63	82	30	95	52	0.5	3.5	0.5
ELG	68	73	24	95	54	0.4	2.5	0.4
ELO	53	68	17	92	28	0.2	1.7	0.7
ERG	74	76	27	96	63	0.5	3.1	0.3

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, RS-Note = Rechtschreibnote, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Für den 1., 2. und 3. Messzeitpunkt und das Kriterium der Rechtschreibnote zum 4. Messzeitpunkt der Form A erreicht die Sensitivität Werte von 44 % bis 74 %, wobei die ELG und ELO-Skala insgesamt am schwächsten und die ERG-Skala Problemschüler am stärksten identifizieren. Die Spezifität erreicht Werte von 68 % bis 89 %, wobei die ELO-Skala insgesamt am schwächsten und die ERG-Skala am stärksten Nicht-Problemschüler identifizieren. Der positive prädiktive Wert liegt zwischen 17 % und 62 %, der negative prädiktive Wert zwischen 82 % und 96 %. Die RAZ-Indexe liegen mit Werten zwischen 25 % und 63 % überwiegend im guten, akzeptablen Bereich, wobei die RAZ-Indexe für die ELO-Skala am schwächsten sind. Die Youden-Indexe erreichen Werte von höchstens Y-I = 0.5. Die positiven und negativen Likelihood-Ratios deuten auf schwache bis kaum relevante diagnostische Evidenz der Form A hin.

In die sechste bis achte klassifikatorische Analyse für die Form A wurden der orthographische Gesamtwert und die Subskalen zum 1., 2. und 3. Messzeitpunkt als Prädiktoren und die Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabellen 45, 46, 47 stellen die Ergebnisse der klassifikatorischen Analysen zur Vorhersage der Ergebnisse im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt dar.

Tabelle 45: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium

Prädiktoren zum 1. MZP (PR ≤ 25)	Kriterium: Rechtschreibfehler im DERET 1-2+ zum 4. MZP (PR ≤ 16)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	79	73	18	98	69	0.5	2.9	0.3
ELG	57	74	15	96	41	0.3	2.2	0.6
ELO	86	72	19	96	79	0.6	3.1	0.2
ERG	71	77	19	97	61	0.5	3.1	0.4

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, , ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 46: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium

Prädiktoren zum 2. MZP (PR ≤ 25)	Kriterium: Rechtschreibfehler im DERET 1-2+ zum 4. MZP (PR ≤ 16)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	65	80	47	90	51	0.5	3.3	0.4
ELG	55	80	43	87	38	0.4	2.8	0.6
ELO	55	80	42	87	38	0.4	2.8	0.6
ERG	74	75	44	91	59	0.5	3.0	0.3

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, , ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 47: Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium

Prädiktoren zum 3. MZP (PR ≤ 25)	Kriterium: Rechtschreibfehler im DERET 1-2+ zum 4. MZP (PR ≤ 16)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	93	80	27	99	90	0.7	4.7	0.1
ELG	71	69	15	97	57	0.4	2.3	0.4
ELO	86	68	17	98	78	0.5	2.7	0.2
ERG	71	74	18	97	60	0.5	2.7	0.4

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, PR = Prozentrang, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Für den 1., 2. und 3. Messzeitpunkt und das Kriterium der Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt der Form A erreicht die Sensitivität Werte von 55 % bis 93 %, wobei die ELG-Skala am schwächsten und der orthographische Gesamtwert und die ELO-Skala insgesamt Problemschüler am stärksten identifizieren. Die Spezifität erreicht Werte von 68 % bis 80 %, wobei keine Skala über alle Messzeitpunkte hinweg am schwächsten oder stärksten Nicht-Problemschüler identifiziert. Der positive prädiktive Wert liegt zwischen 15 % und 47 %, der negative prädiktive Wert zwischen 87 % und 99 %. Die RAZ-Indexe liegen mit Werten zwischen 38 % und 90 % im guten bis sehr guten Bereich, wobei die RAZ-Indexe des orthographischen Gesamtwerts und der ELO-Skala zum 1. und 3. Messzeitpunkt am höchsten sind. Die Youden-Indexe erreichen Werte von höchstens Y-I = 0.7. Die positiven und negativen Likelihood-Ratios deuten insgesamt auf schwache bis kaum relevante diagnostische Evidenz der Form A hin.

5.2.3.5.2 Prognostische Validität der Form B

Nachfolgend werden die klassifikatorischen Analysen für die FE-RS 2 der Form B dargestellt. In die erste klassifikatorische Analyse wurden der orthographische Gesamtwert und die Subskalen zum 1. Messzeitpunkt als Prädiktoren und der orthographische Gesamtwert zum 3. Messzeitpunkt als Kriterium aufgenommen. In die zweite

klassifikatorische Analyse wurden der orthographische Gesamtwert und die Subskalen zum 2. Messzeitpunkt als Prädiktoren und der orthographische Gesamtwert zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabellen 48 und 49 stellen die Ergebnisse der klassifikatorischen Analysen zur Vorhersage des orthographischen Gesamtwerts der Form B zum 3. und 4. Messzeitpunkt dar.

Tabelle 48: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 1. Messzeitpunkt und des Kriteriums zum 3. Messzeitpunkt

Prädiktoren zum 1. MZP (PR ≤ 25)	Kriterium: Orthographischer Gesamtwert zum 3. MZP (PR ≤ 15)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	77	80	47	94	67	0.6	3.9	0.3
ELG	67	79	42	91	54	0.5	3.2	0.4
ELO	61	75	35	89	42	0.4	2.4	0.5
ERG	67	76	39	91	52	0.4	2.8	0.4

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 49: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 2. Messzeitpunkt und des Kriteriums zum 4. Messzeitpunkt

Prädiktoren zum 2. MZP (PR ≤ 25)	Kriterium: Orthographischer Gesamtwert zum 4. MZP (PR ≤ 15)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	73	84	52	93	63	0.6	4.6	0.3
ELG	68	77	41	91	53	0.5	3.0	0.4
ELO	78	79	47	94	68	0.6	3.7	0.3
ERG	70	78	43	92	57	0.5	3.2	0.4

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Für den 1. und 2. Messzeitpunkt und das Kriterium des orthographischen Gesamtwerts zum 4. Messzeitpunkt der Form B erreicht die Sensitivität Werte von 61 % bis 77 %, wobei zum 1. Messzeitpunkt der orthographische Gesamtwert und zum 2. Messzeitpunkt

die ELO-Skala Problemschüler am stärksten identifizieren. Die Spezifität erreicht Werte von 75 % bis 84 %, wobei der orthographische Gesamtwert am stärksten Nicht-Problemschüler identifiziert. Der positive prädiktive Wert liegt zwischen 35 % und 52 %, der negative prädiktive Wert zwischen 89 % und 94 %. Die RAZ-Indexe liegen mit Werten zwischen 42 % und 68 % überwiegend im guten, akzeptablen Bereich, wobei zum 1. Messzeitpunkt die RAZ-Indexe des orthographischen Gesamtwerts und zum 2. Messzeitpunkt die der ELO-Skala am höchsten sind. Die Youden-Indexe erreichen Werte von höchstens $Y-I = 0.6$. Die positiven und negativen Likelihood-Ratios deuten auf schwache diagnostische Evidenz der Form B hin.

In die dritte bis fünfte klassifikatorische Analyse für die Form B wurden der orthographische Gesamtwert und die Subskalen zum 1., 2. und 3. Messzeitpunkt als Prädiktoren und die Rechtschreibnote zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabellen 50, 51 und 52 stellen die Ergebnisse der klassifikatorischen Analysen zur Vorhersage der Rechtschreibnote zum 4. Messzeitpunkt dar.

Tabelle 50: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt

Prädiktoren zum 1. MZP (PR \leq 25)	Kriterium: Rechtschreibnote zum 4. MZP (RS-Note \geq 4)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	68	81	52	89	54	0.5	3.6	0.4
ELG	63	83	53	88	49	0.5	3.7	0.4
ELO	49	78	40	83	28	0.3	2.2	0.7
ERG	66	75	45	88	48	0.4	2.6	0.5

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, RS-Note = Rechtschreibnote, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 51: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt

Prädiktoren zum 2. MZP (PR \leq 25)	Kriterium: Rechtschreibnote zum 4. MZP (RS-Note \geq 4)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	53	79	23	93	37	0.3	2.5	0.6
ELG	68	75	25	95	55	0.4	2.7	0.4
ELO	53	71	28	93	31	0.2	1.8	0.7
ERG	68	73	24	95	54	0.4	2.5	0.4

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, RS-Note = Rechtschreibnote, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 52: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt

Prädiktoren zum 3. MZP (PR \leq 25)	Kriterium: Rechtschreibnote zum 4. MZP (RS-Note \geq 4)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	63	84	55	88	50	0.5	3.9	0.4
ELG	56	84	51	86	41	0.4	3.5	0.5
ELO	46	85	49	84	33	0.3	3.1	0.6
ERG	78	72	46	92	64	0.5	2.8	0.3

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, RS-Note = Rechtschreibnote, SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Für den 1., 2. und 3. Messzeitpunkt und das Kriterium der Rechtschreibnote zum 4. Messzeitpunkt der Form B erreicht die Sensitivität Werte von 46 % bis 78 %, wobei die ELO-Skala am schwächsten und die ERG-Skala Problemschüler am stärksten identifizieren. Die Spezifität erreicht Werte von 71 % bis 85 %, wobei keine Skala über alle Messzeitpunkte hinweg am schwächsten oder stärksten Nicht-Problemschüler identifiziert. Der positive prädiktive Wert liegt zwischen 23 % und 55 %, der negative prädiktive Wert zwischen 83 % und 95 %. Die RAZ-Inde lie mit Werten zwischen 28 % und 64 % überwiegend im guten, akzeptablen Bereich, wobei die RAZ-Inde für die ELO-Skala am schwächsten sind. Die Youden-Inde erreichen Werte von höchstens

Y-I = 0.5. Die positiven und negativen Likelihood-Ratios deuten auf schwache bis kaum relevante diagnostische Evidenz der Form B hin.

In die sechste bis achte klassifikatorische Analyse für die Form B wurden der orthographische Gesamtwert und die Subskalen zum 1., 2. und 3. Messzeitpunkt als Prädiktoren und die Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt als Kriterium aufgenommen. Tabellen 53, 54, 55 stellen die Ergebnisse der klassifikatorischen Analysen zur Vorhersage der Ergebnisse im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt dar.

Tabelle 53: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium

Prädiktoren zum 1. MZP (PR ≤ 25)	Kriterium: Rechtschreibfehler im DERET 1-2+ zum 4. MZP (PR ≤ 16)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	71	80	49	91	59	0.5	3.6	0.4
ELG	61	79	44	88	45	0.4	2.9	0.5
ELO	59	75	39	87	40	0.3	2.4	0.5
ERG	63	76	41	89	46	0.4	2.6	0.5

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 54: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium

Prädiktoren zum 2. MZP (PR ≤ 25)	Kriterium: Rechtschreibfehler im DERET 1-2+ zum 4. MZP (PR ≤ 16)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	93	78	25	99	90	0.7	4.2	0.1
ELG	71	72	16	97	58	0.4	2.5	0.4
ELO	71	71	16	97	58	0.4	2.4	0.4
ERG	71	72	16	97	58	0.4	2.5	0.4

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Tabelle 55: Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium

Prädiktoren zum 3. MZP (PR ≤ 25)	Kriterium: Rechtschreibfehler im DERET 1-2+ zum 4. MZP (PR ≤ 16)							
	SEN	SPE	PPW	NPW	R-I	Y-I	LR+	LR-
OG	76	81	52	93	65	0.6	4.0	0.3
ELG	63	80	46	89	48	0.4	3.2	0.5
ELO	51	81	42	86	34	0.3	2.7	0.6
ERG	88	71	45	96	79	0.6	3.0	0.2

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, PR = Prozentrang, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), SEN = Sensitivität in Prozent, SPE = Spezifität in Prozent, PPW = positiver prädiktiver Wert in Prozent, NPW = negativer prädiktiver Wert in Prozent, R-I = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent, Y-I = Youden-Index, LR+ = positiver Likelihood-Ratio, LR- = negativer Likelihood-Ratio, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Für den 1., 2. und 3. Messzeitpunkt und das Kriterium der Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt der Form B erreicht die Sensitivität Werte von 51 % bis 93 %, wobei die ELO-Skala insgesamt am schwächsten und der orthographische Gesamtwert zum 1. und 2. Messzeitpunkt Problemschüler am stärksten identifizieren. Die Spezifität erreicht Werte von 71 % bis 81 %, wobei der orthographische Gesamtwert insgesamt am stärksten Nicht-Problemschüler identifiziert. Der positive prädiktive Wert liegt zwischen 16 % und 52 %, der negative prädiktive Wert zwischen 86 % und 99 %. Die RAZ-Inde liegen mit Werten zwischen

34 % und 90 % überwiegend im guten, akzeptablen Bereich. Die Youden-Indexe erreichen Werte von höchstens $Y-I = 0.7$. Die positiven und negativen Likelihood-Ratios deuten, bis auf den orthographischen Gesamtwert zum 2. Messzeitpunkt, insgesamt auf schwache diagnostische Evidenz der Form B hin.

Zusammenfassend betrachtet, ist die Vorhersageleistung der Rechtschreibleistung zum 3. Messzeitpunkt durch den orthographischen Gesamtwert des 1. Messzeitpunkts für beide Formen der FE-RS 2 mit RATZ-Indexen von 69 % und 67 % annähernd gleich hoch. Für die Vorhersage der Rechtschreibleistung zum 4. Messzeitpunkt durch den orthographischen Gesamtwert zum 2. Messzeitpunkt leistet die Form A mit einem RATZ-Index von 72 % einen höheren Beitrag als die Form B. Eine frühe und sichere Vorhersage der Rechtschreibleistung ist somit am Ende des ersten Schulhalbjahres der 2. Klassenstufe durch den orthographischen Gesamtwert der Form A möglich. Damit bestätigen die klassifikatorischen Analysen die Ergebnisse der Regressionsanalysen. Die Vorhersageleistung der Rechtschreibleistung durch die Subskalen ist fast über alle Messzeitpunkte nicht spezifisch. Die Vorhersage der Rechtschreibnote am Ende der 2. Klassenstufe ist ebenfalls unzureichend, da beide Formen der FE-RS 2 RATZ-Indexe von höchstens 63 % bzw. 64 % erreichen. Die Vorhersage der Anzahl der Rechtschreibfehler im DERET 1-2+ (Stock & Schneider, 2008a) am Ende der 2. Klassenstufe ist mit beiden Formen der FE-RS 2 möglich. Für die Form A sind dazu der orthographische Gesamtwert und die ELO-Subskala zum 1. und 3. Messzeitpunkt geeignet. Für die Form B sind der orthographische Gesamtwert zum 2. Messzeitpunkt und die ERG-Subskala zum 3. Messzeitpunkt geeignet.

H_{4,1}: Der orthographische Gesamtwert und die Subskalen der Form A und der Form B zum 1., 2. und 3. Messzeitpunkt tragen signifikant zur Varianzaufklärung des orthographischen Gesamtwerts zum 2., 3. und 4. Messzeitpunkt bei. Für den RATZ-Index gilt $> 66\%$.

Der orthographische Gesamtwert für beide Formen der FE-RS 2 trägt zu allen Messzeitpunkten signifikant zur Varianzaufklärung des orthographischen Gesamtwerts späterer Messzeitpunkte bei. Eine relativ frühe und sichere Vorhersage der Rechtschreibleistung am Ende der 2. Klassenstufe ist mit der FE-RS 2 der Form A zum 2. Messzeitpunkt, das heißt zum Ende des ersten Schulhalbjahres, möglich. Die

klassifikatorischen Analysen bestätigen dieses Ergebnis, denn die RATZ-Indexe sind für den orthographischen Gesamtwert der Form A etwas höher als für die Form B. Damit ist der orthographische Gesamtwert besser als die Subskalen für die Vorhersage der Rechtschreibleistungen geeignet. Die Hypothese $H_{4.1}$ lässt sich deshalb nur zum Teil bestätigen.

H_{4.2}: Der orthographische Gesamtwert und die Subskalen der Form A und der Form B zum 1., 2. und 3. Messzeitpunkt sagen die Rechtschreibnote zum 4. Messzeitpunkt vorher. Für den RATZ-Index gilt > 66 %.

Für die Vorhersage der Rechtschreibnote zum 4. Messzeitpunkt durch den orthographischen Gesamtwert und die Subskalen beider Formen der FE-RS 2 wird kein RATZ-Index von > 66 % erreicht. Die Hypothese $H_{4.2}$ lässt sich deshalb nicht bestätigen. Die Folge ist, dass die Vorhersageleistung beider Formen der FE-RS 2 für die Rechtschreibnote am Schuljahresende der 2. Klassenstufe nur unspezifisch ist.

5.2.4 Äquivalenz der Parallelversionen

Für die Überprüfung der Äquivalenzkriterien wurden die Varianzen, Mittelwerte, Häufigkeitsverteilungen sowie die Reliabilitäts- und Validitätskorrelationen der FE-RS 2 der Form A und B miteinander verglichen.

Die Überprüfung der Varianzhomogenität der FE-RS 2 erfolgte mit dem Levene-Test. Der Levene-Test liefert überwiegend signifikante Ergebnisse. Das bedeutet, dass die Varianzen der Formen A und B überwiegend verschieden sind. Der Vergleich der Mittelwerte der Formen A und B wurde aufgrund der Varianzheterogenität und der nicht normalverteilten Merkmalsausprägungen mit dem U-Test durchgeführt. Die Mittelwertvergleiche liefern überwiegend signifikante Ergebnisse. Das bedeutet, dass die Mittelwerte der Formen A und B überwiegend statistisch bedeutsam verschieden sind. Tabelle 56 gibt die Ergebnisse der Äquivalenzüberprüfung von Varianzen und Mittelwerten der FE-RS 2 wieder.

Tabelle 56: Äquivalenzüberprüfung von Varianzen und Mittelwerten der FE-RS 2

	Varianzvergleich	Mittelwertvergleich
	p	p
OG 1. MZP	.05 [*]	.00 ^{***}
OG 2. MZP	.09 ^{ns}	.01 ^{**}
OG 3. MZP	.39 ^{ns}	.04 [*]
OG 4. MZP	.13 ^{ns}	.02 [*]
ELG 1. MZP	.55 ^{ns}	.00 ^{***}
ELG 2. MZP	.00 ^{***}	.00 ^{***}
ELG 3. MZP	.03 [*]	.00 ^{***}
ELG 4. MZP	.05 [*]	.00 ^{***}
ELO 1. MZP	.08 ^{ns}	.98 ^{ns}
ELO 2. MZP	.27 ^{ns}	.00 ^{***}
ELO 3. MZP	.11 ^{ns}	.69 ^{ns}
ELO 4. MZP	.03 [*]	.00 ^{***}
ERG 1. MZP	.01 ^{**}	.33 ^{ns}
ERG 2. MZP	.35 ^{ns}	.46 ^{ns}
ERG 3. MZP	.48 ^{ns}	.01 ^{**}
ERG 4. MZP	.03 [*]	.21 ^{ns}

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, MZP = Messzeitpunkt, p = Irrtumswahrscheinlichkeit, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$

H_{5.1}: Die Varianzen und Mittelwerte zwischen den orthographischen Gesamtwerten, den Subskalen der Form A und Form B zu den 4 Messzeitpunkten unterscheiden sich nicht signifikant.

Die Varianzen und Mittelwerte zwischen den orthographischen Gesamtwerten, den Subskalen der Form A und Form B zu den 4 Messzeitpunkten unterscheiden sich überwiegend signifikant. Die Hypothese H_{5.1} lässt sich deshalb nicht bestätigen.

Zur Überprüfung der Äquivalenz der Häufigkeitsverteilungen wurden die Häufigkeitspolygone der FE-RS 2 der Form A und B miteinander verglichen. Die Häufigkeitspolygone für den orthographischen Gesamtwert und die Subskalen beider Formen sind über alle 4 Messzeitpunkte nicht deckungsgleich. Sie unterscheiden sich augenscheinlich in ihrer exzessiven Verteilung und in ihrer Gipfelhöhe.

H_{5.2}: Die Häufigkeitsverteilungen der Form A und Form B unterscheiden sich nicht voneinander.

Die Häufigkeitsverteilungen der Form A und Form B unterscheiden sich augenscheinlich voneinander. Die Hypothese $H_{5.2}$ lässt sich deshalb nicht bestätigen.

Die Überprüfung der Äquivalenz der Reliabilität erfolgte durch die Berechnung der Paralleltestreliabilität und durch die Berechnung der Prüfgröße z (Bortz & Schuster, 2010). Für die Paralleltestreliabilitäten des orthographischen Gesamtwerts und der Subskalen der FE-RS 2 wurden die Korrelationskoeffizienten nach Pearson und Spearman zum 1. und 4. Messzeitpunkt berechnet. Für den orthographischen Gesamtwert fallen die Korrelationen am höchsten aus; sie liegen mit $r = .70$ und $r = .83$ im mittleren und hohen Bereich. Die Korrelationskoeffizienten für die Subskalen verweisen zum 1. und 4. Messzeitpunkt auf niedrige bis mittlere Korrelationen. Zusammenfassend betrachtet, ist die gemittelte Paralleltestreliabilität der FE-RS 2 mittelhoch. Das heißt, die Form A und Form B messen im mittleren Maße gleich genau. Tabelle 57 gibt die Paralleltestkoeffizienten für den orthographischen Gesamtwert und die Subskalen der FE-RS 2 zum 1. und 4. Messzeitpunkt wieder.

Tabelle 57: Paralleltestreliabilität der FE-RS 2 zum 1. Messzeitpunkt ($N = 68$) und 4. Messzeitpunkt ($N = 63$)

	OG _{B1}	ELG _{B1}	ELO _{B1}	ERG _{B1}	OG _{B4}	ELG _{B4}	ELO _{B4}	ERG _{B4}
OG _{A1}	.70 ^{** (P)}							
ELG _{A1}		.67 ^{** (S)}						
ELO _{A1}			.55 ^{** (S)}					
ERG _{A1}				.56 ^{** (S)}				
OG _{A4}					.83 ^{** (S)}			
ELG _{A4}						.49 ^{** (S)}		
ELO _{A4}							.55 ^{** (S)}	
ERG _{A4}								.65 ^{** (S)}

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, A₁ = Form A zum 1. Messzeitpunkt, A₂ = Form A zum 2. Messzeitpunkt, A₃ = Form A zum 3. Messzeitpunkt, A₄ = Form A zum 4. Messzeitpunkt, B₁ = Form B zum 1. Messzeitpunkt, B₂ = Form B zum 2. Messzeitpunkt, B₃ = Form B zum 3. Messzeitpunkt, B₄ = Form B zum 4. Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, ERO = erweitert orthographisches Können, N = Anzahl, (P) = Korrelationskoeffizient nach Pearson, (S) = Korrelationskoeffizient nach Spearman, ** = $p \leq 0.01$ (2-seitig)

H_{5.3}: Für die Paralleltestreliabilität des orthographischen Gesamtwerts und der Subskalen der Form A und Form B gilt $r_u > .70$.

Die Werte für die Paralleltestreliabilität des orthographischen Gesamtwerts und der

Subskalen der FE-RS 2 liegen zum 1. und 4. Messzeitpunkt, bis auf den orthographischen Gesamtwert zum 4. Messzeitpunkt, im mittleren Bereich. Die Paralleltestreliabilität des orthographischen Gesamtwerts zum 4. Messzeitpunkt liegt im hohen Bereich. Die Hypothese $H_{5.3}$ lässt sich deshalb nicht eindeutig bestätigen. Für den orthographischen Gesamtwert der FE-RS 2 kann zu beiden Messzeitpunkten von Parallelität ausgegangen werden, da der Wert zum 1. Messzeitpunkt mit $r = .70$ nahe am hohen Bereich liegt. Für die Subskalen kann nicht von Parallelität ausgegangen werden, da alle Werte $r < .70$ sind. Somit ist die Paralleltestreliabilität der FE-RS 2 nur für den orthographischen Gesamtwert zufriedenstellend hoch.

Die Prüfgröße z wurde zur Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Reliabilitätskoeffizienten herangezogen. Es wurde eine zweiseitige Irrtumswahrscheinlichkeit von 5 % angenommen, sodass sich kritische Grenzwerte von $z_{\text{krit}} = 1.96$ und $z_{\text{krit}} = -1.96$ ergeben. Befindet sich der empirisch ermittelte z -Wert innerhalb der kritischen Grenzwerte, bedeutet das, dass sich die Korrelationen nicht signifikant voneinander unterscheiden (Bortz & Schuster, 2010).

Für die Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Retest-Koeffizienten der FE-RS 2 der Form A und der Form B nehmen alle ermittelten z -Werte Größen innerhalb der kritischen Grenzwerte an. Das bedeutet, dass die Form A und Form B hinsichtlich der Retestreliabilität gleich korrelieren. Tabelle 58 stellt die ermittelten Prüfgrößen dar. Zur besseren Vergleichbarkeit werden die Retest-Koeffizienten beider Formen wiederholt aufgeführt.

Tabelle 58: Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Retest-Koeffizienten der Form A und B

Form A	r_{tt}	Form B	r_{tt}	z
OG _{A1→A3}	.72**	OG _{B1/B3}	.71**	0.21
OG _{A2→A4}	.79**	OG _{B2/B4}	.72**	1.66
ELG _{A1→A3}	.55**	ELG _{B1/B3}	.64**	-1.43
ELG _{A2→A4}	.62**	ELG _{B2/B4}	.51**	1.65
ELO _{A1→A3}	.48**	ELO _{B1/B3}	.56**	-1.12
ELO _{A2→A4}	.47**	ELO _{B2/B4}	.49**	-0.26
ERG _{A1→A3}	.61**	ERG _{B1/B3}	.55**	0.93
ERG _{A2→A4}	.67**	ERG _{B2/B4}	.59**	1.35

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, A₁ = Form A zum 1. Messzeitpunkt, A₂ = Form A zum 2. Messzeitpunkt, A₃ = Form A zum 3. Messzeitpunkt, A₄ = Form A zum 4. Messzeitpunkt, B₁ = Form B zum 1. Messzeitpunkt, B₂ = Form B zum 2. Messzeitpunkt, B₃ = Form B zum 3. Messzeitpunkt, B₄ = Form B zum 4. Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, r_{tt} = Retest-Koeffizient, z = Prüfgröße, ** = $p \leq 0.01$ (2-seitig)

Für die Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Testhalbierungs-Koeffizienten der Form A und der Form B nehmen alle ermittelten z-Werte, bis auf die ermittelten Werte zwischen ELG_{A3} und ELG_{B3}, ERG_{A1} und ERG_{B1} sowie ERG_{A4} und ERG_{B4}, Größen innerhalb der kritischen Grenzwerte an. Das bedeutet, dass die Form A und Form B hinsichtlich der Testhalbierungsreliabilität überwiegend gleich korrelieren. Tabelle 59 stellt die ermittelten Prüfgrößen dar. Zur besseren Vergleichbarkeit werden die Testhalbierungs-Koeffizienten beider Formen wiederholt aufgeführt.

Tabelle 59: Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Testhalbierungs-Koeffizienten der Form A und B

Form A	r_{tt}	Form B	r_{tt}	z
OG 1. MZP	.87	OG 1. MZP	.89	-0.91
OG 2. MZP	.87	OG 2. MZP	.87	0
OG 3. MZP	.86	OG 3. MZP	.85	0.38
OG 4. MZP	.84	OG 4. MZP	.86	-0.73
ELG 1. MZP	.68	ELG 1. MZP	.73	-1.02
ELG 2. MZP	.78	ELG 2. MZP	.74	0.97
ELG 3. MZP	.71	ELG 3. MZP	.57	2.43 ^(s.)
ELG 4. MZP	.65	ELG 4. MZP	.71	-1.14
ELO 1. MZP	.72	ELO 1. MZP	.74	-0.43
ELO 2. MZP	.79	ELO 2. MZP	.82	-0.88
ELO 3. MZP	.74	ELO 3. MZP	.74	0
ELO 4. MZP	.64	ELO 4. MZP	.65	-0.17
ERG 1. MZP	.64	ERG 1. MZP	.80	-3.47 ^(s.)
ERG 2. MZP	.76	ERG 2. MZP	.73	0.68
ERG 3. MZP	.60	ERG 3. MZP	.67	-1.20
ERG 4. MZP	.74	ERG 4. MZP	.61	2.45 ^(s.)

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, r_{tt} = Testhalbierungskoeffizient, z = Prüfgröße, s. = signifikanter Korrelationsunterschied

Für die Prüfung der Signifikanz von Korrelationsunterschieden zwischen den internen Konsistenzen (Alphakoeffizienten) der Form A und der Form B nehmen alle ermittelten z -Werte Größen innerhalb der kritischen Grenzwerte an. Das bedeutet, dass die Form A und Form B hinsichtlich der internen Konsistenzen gleich korrelieren. Tabelle 60 stellt die ermittelten Prüfgrößen dar. Zur besseren Vergleichbarkeit werden die internen Konsistenzen beider Formen wiederholt aufgeführt.

Tabelle 60: Prüfung der Signifikanz von Korrelationsunterschieden zwischen den internen Konsistenzen der Form A und B

Form A	interne Konsistenz	Form B	interne Konsistenz	z
OG 1. MZP	.84	OG 1. MZP	.88	-1.58
OG 2. MZP	.87	OG 2. MZP	.87	0
OG 3. MZP	.86	OG 3. MZP	.86	0
OG 4. MZP	.85	OG 4. MZP	.85	0
ELG 1. MZP	.70	ELG 1. MZP	.74	-0.85
ELG 2. MZP	.77	ELG 2. MZP	.72	1.14
ELG 3. MZP	.72	ELG 3. MZP	.64	1.53
ELG 4. MZP	.61	ELG 4. MZP	.68	-1.22
ELO 1. MZP	.68	ELO 1. MZP	.75	-1.47
ELO 2. MZP	.73	ELO 2. MZP	.75	-0.45
ELO 3. MZP	.69	ELO 3. MZP	.68	0.19
ELO 4. MZP	.68	ELO 4. MZP	.69	-0.19
ERG 1. MZP	.65	ERG 1. MZP	.73	-1.57
ERG 2. MZP	.74	ERG 2. MZP	.67	1.42
ERG 3. MZP	.62	ERG 3. MZP	.64	-0.34
ERG 4. MZP	.67	ERG 4. MZP	.57	1.66

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, z = Prüfgröße, MZP = Messzeitpunkt

H_{5.4}: Die Reliabilitätskoeffizienten der Form A und Form B unterscheiden sich nicht signifikant.

Zusammenfassend betrachtet, unterscheiden sich die Reliabilitätskoeffizienten der Form A und Form B, bis auf wenige Ausnahmen bezüglich der Werte der Testhalbierungsreliabilität, nicht signifikant. Die Hypothese H_{5.4} lässt sich deshalb bestätigen.

Für die Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Validitätskoeffizienten der Form A und B nehmen alle ermittelten z-Werte, bis auf die ermittelten Werte zwischen dem orthographischen Gesamtwert der FE-RS 2 und des FE-L 2 (Kuhlmann & Hartke, 2009b) zum 1. und 4. Messzeitpunkt sowie zwischen dem orthographischen Gesamtwert der FE-RS 2 und des DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt, Größen innerhalb der kritischen Grenzwerte an. Das bedeutet, dass die Form A und B hinsichtlich der Validität überwiegend gleich korrelieren. Tabelle 61 stellt die ermittelten Prüfgrößen dar. Zur besseren Vergleichbarkeit werden die Validitätskoeffizienten beider Formen wiederholt aufgeführt.

Tabelle 61: Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Validitätskoeffizienten der Form A und B

Form A	r	Form B	r	z
1. MZP				
DERET 1-2+	-.71**	DERET 1-2+	-.77**	-1.06
FE-L 2	.47**	FE-L 2	.68**	-3.25 ^(s.)
FE-AF 2	-.1	FE-AF 2	.07	0.29
DEMAT 1+	-.02	DEMAT 1+	-.06	-0.29
4. MZP				
DERET 1-2+	-.81**	DERET 1-2+	-.69**	2.84 ^(s.)
FE-L 2	.68**	FE-L 2	.50**	2.85 ^(s.)
FE-AF 2	.05	FE-AF 2	.02	0.29
DEMAT 2+	-.02	DEMAT 2+	.03	-0.10
Rechtschreibnote	-.72**	Rechtschreibnote	-.62**	1.70

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), FE-L 2 = Formative Erfassung der Lesefertigkeit im 2. Schuljahr (Kuhlmann & Hartke, 2009b), FE-AF 2 = Formative Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr (Kuhlmann & Hartke, 2009a), DEMAT 1+ = Deutscher Mathematiktest für erste Klassen (Krajewski et al., 2002), DEMAT 2+ = Deutscher Mathematiktest für zweite Klassen (Krajewski et al., 2004), MZP = Messzeitpunkt, r = Korrelationskoeffizient, z = Prüfgröße, s. = signifikanter Korrelationsunterschied, ** = $p \leq 0.01$ (2-seitig)

H_{5.5}: Die Validitätskoeffizienten der Form A und Form B unterscheiden sich nicht signifikant.

Die Validitätskoeffizienten der Form A und Form B unterscheiden sich, bis auf wenige Ausnahmen, nicht signifikant. Die Hypothese H_{5.5} lässt sich deshalb bestätigen.

Zusammenfassend betrachtet, zeigt sich, dass die Formen A und B der FE-RS 2 überwiegend nicht äquivalent sind. Die beiden Formen der FE-RS 2 unterscheiden sich in den Häufigkeitsverteilungen, den Varianzen und den Mittelwerten. Die Ungleichheit der Häufigkeitsverteilungen ist meistens „auf eine ungleichwertige Anordnung der Aufgaben hinsichtlich ihrer Schwierigkeit zurückzuführen“ (Lienert & Raatz, 1998, S. 301). Die Ungleichheit der Varianzen betrifft vor allem die Subskalen der FE-RS 2, denn für den 2., 3. und 4. Messzeitpunkt sind die Varianzen des orthographischen Gesamtwerts gleich. Die Mittelwerte sind über beide Formen und alle Messzeitpunkte, bis auf 5 Ausnahmen, signifikant verschieden.

5.2.5 Veränderungssensitivität

In diesem Abschnitt werden die Ergebnisse zur Veränderungssensitivität für den orthographischen Gesamtwert und die Subskalen der FE-RS 2 dargelegt. Zuerst werden die Verteilungen des orthographischen Gesamtwerts und der Subskalen beschrieben. Anschließend werden die Ergebnisse der Mittelwertvergleiche sowie die Häufigkeitsverteilungen und Kennwerte für die Mittelwertdifferenzen dargelegt. Die Prüfung auf Veränderungssensitivität erfolgte aufgrund des Untersuchungsdesigns zwischen dem 1. und 3. sowie zwischen dem 2. und 4. Messzeitpunkt.

5.2.5.1 Häufigkeitsveränderungen und Mittelwertdifferenzen der Form A

Die Häufigkeitsverteilungen für den orthographischen Gesamtwert und die Subskalen der Form A zwischen dem 1. und 3. sowie zwischen dem 2. und 4. Messzeitpunkt sind in den Abbildungen 26, 27, 28 und 29 dargestellt.

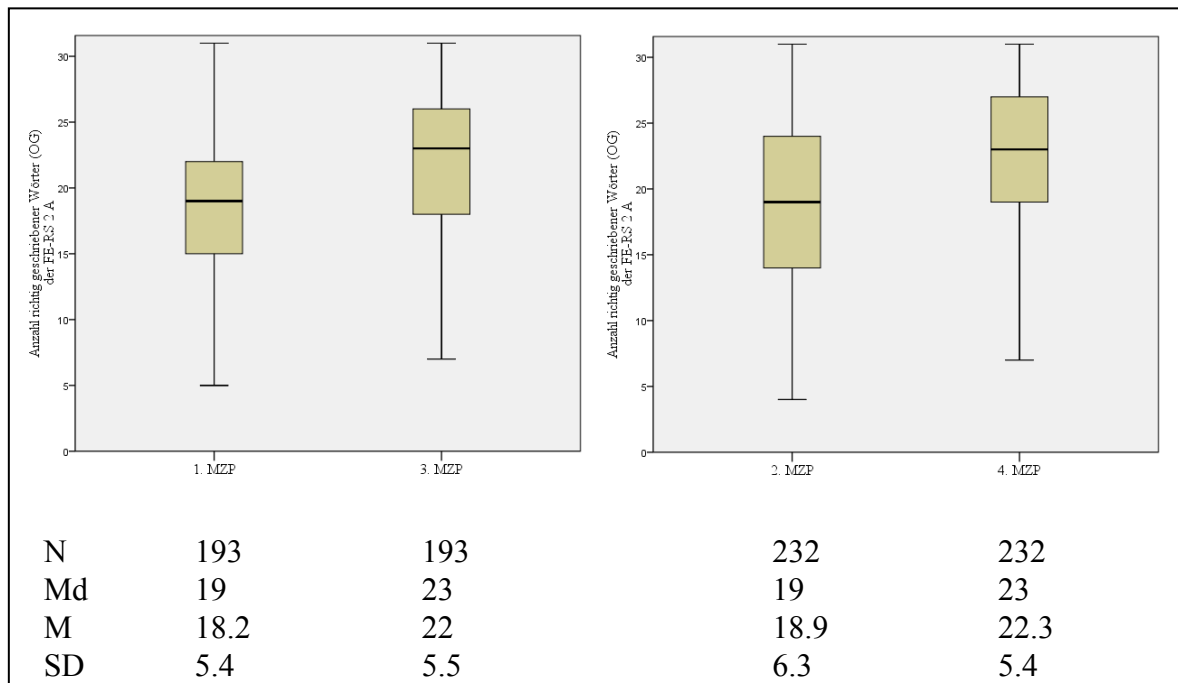


Abbildung 26: Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, OG = orthographischer Gesamtwert, N = Anzahl, Md = Median, M = Mittelwert, SD = Standardabweichung

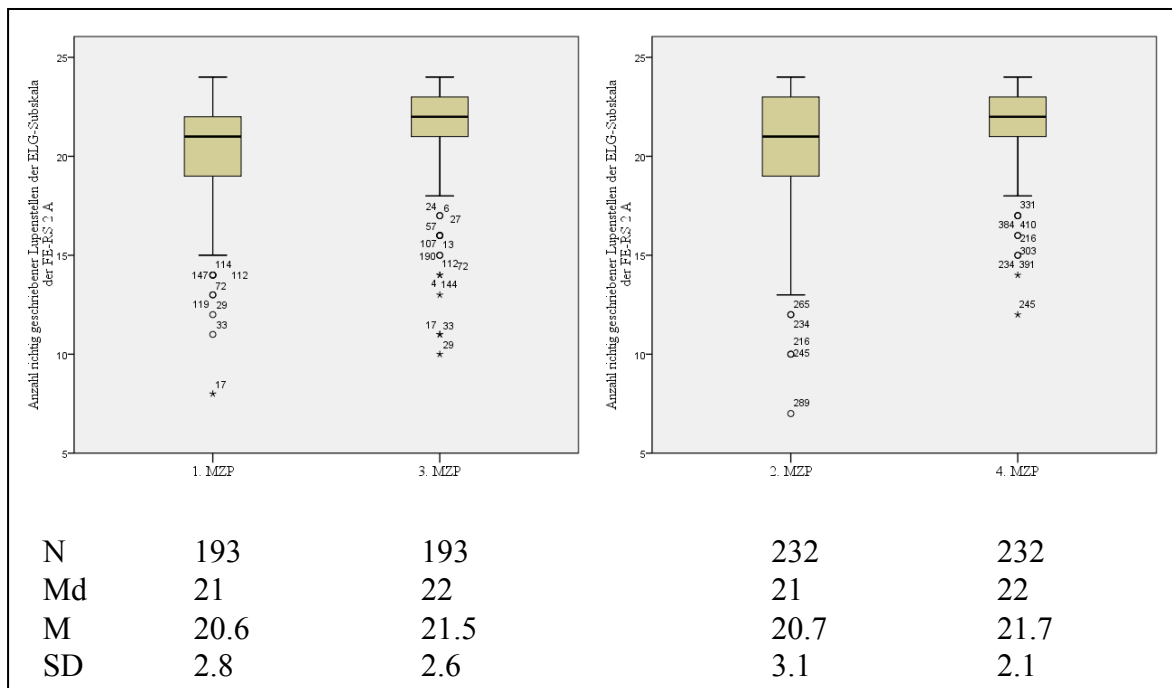


Abbildung 27: Häufigkeitsverteilungen für die ELG-Subskala der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELG = elementar graphematisches Können, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

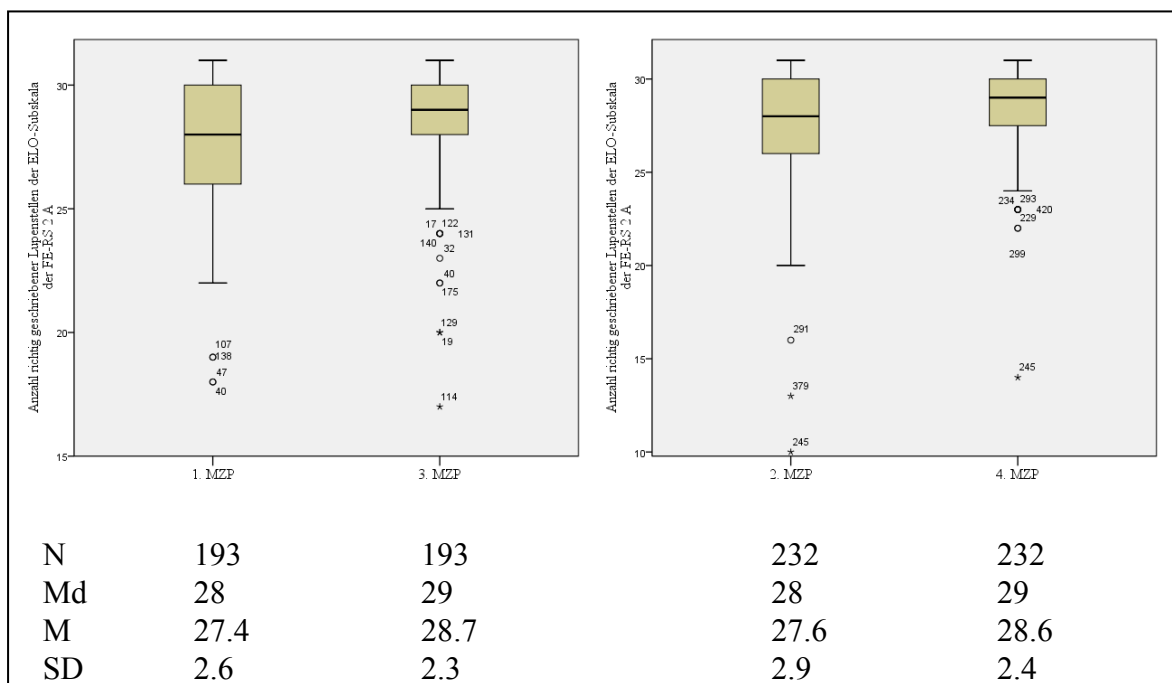


Abbildung 28: Häufigkeitsverteilungen für die ELO-Subskala der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELO = elementar orthographisches Können, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

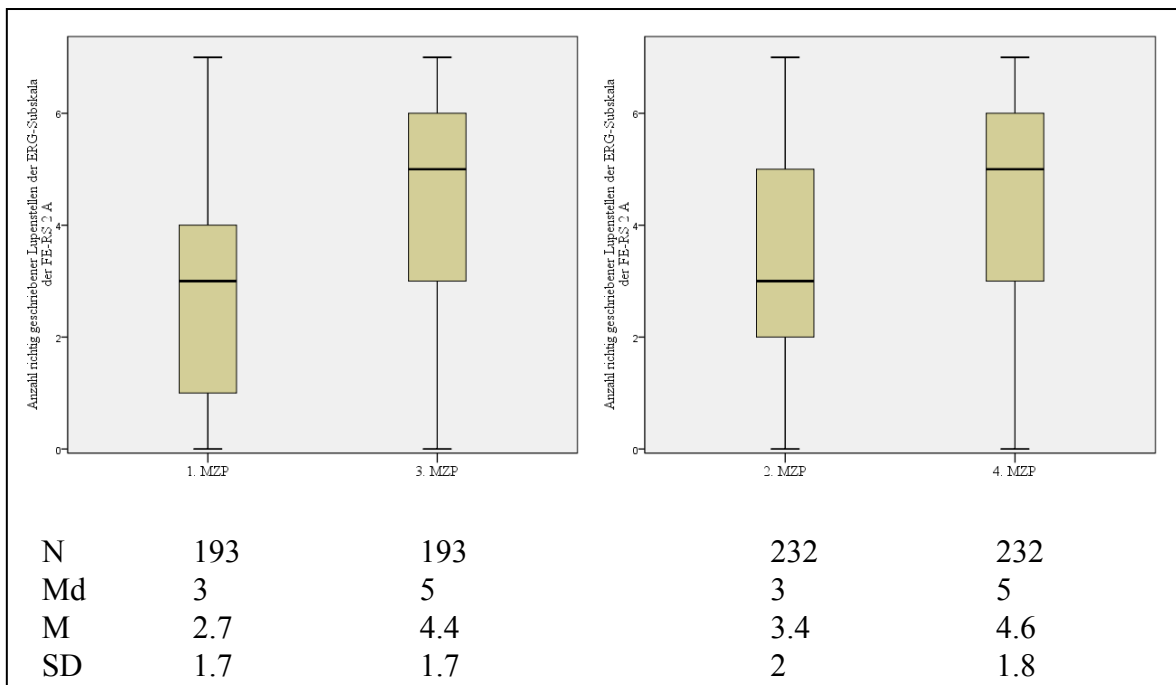


Abbildung 29: Häufigkeitsverteilungen für die ERG-Subskala der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ERG = erweitert graphematisches Können, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

Wie in den Abbildungen 26 bis 29 zu erkennen, überschneiden sich die Interquartilbereiche der Form A, vor allem der ELG- und ELO-Subskalen, sehr stark. Auch die Mediane und Mittelwerte, vor allem der ELG- und ELO-Subskalen, unterscheiden sich kaum zwischen den Messzeitpunkten. Somit sind augenscheinlich deutliche Veränderungen nur bei dem orthographischen Gesamtwert und der ERG-Subskala erkennbar. Die Standardabweichungen des orthographischen Gesamtwerts und der Subskalen bleiben vom 1. zum 3. Messzeitpunkt nahezu konstant, verkleinern sich aber vom 2. zum 4. Messzeitpunkt. Das bedeutet, dass die Merkmalsausprägungen vom 1. zum 3. Messzeitpunkt nahezu gleich stark streuen und die Stichprobe hinsichtlich der Rechtschreibleistungen relativ heterogen ist. Vom 2. zum 4. Messzeitpunkt streuen die Merkmalsausprägungen weniger stark. Die Stichprobe wird hinsichtlich der Rechtschreibleistungen etwas homogener.

Der Vergleich der Mittelwerte der Form A wurde mit dem Wilcoxon-Test durchgeführt, da keine Normalverteilungen vorlagen. Die Ergebnisse des Mittelwertvergleichs zeigen höchst signifikante Unterschiede zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt des orthographischen Gesamtwerts und der Subskalen. Vom 1. zum 3. sowie vom 2. zum 4. Messzeitpunkt steigen für den orthographischen Gesamtwert und die

Subskalen der Form A die Anzahlen richtig geschriebener Wörter und Lupenstellen in mindestens 55 % der Fälle. Damit ist in diesen Fällen ein Lernfortschritt im Rechtschreiben zu erkennen. In maximal 28 % der Fälle sinkt und in maximal 23 % der Fälle stagniert die Anzahl richtig geschriebener Wörter und Lupenstellen. In diesen Fällen ist kein Lernfortschritt im Rechtschreiben zu erkennen. Alle einzelnen positiven, negativen und Null differenzen sind in Tabelle 62 zusammengefasst.

Tabelle 62: Mittelwertvergleich des orthographischen Gesamtwerts und der Subskalen der Form A

	N	Positive Differenzen	Negative Differenzen	Null- differenzen	p
OG 3. MZP - OG 1. MZP	193	155	25	13	0.000
OG 4. MZP - OG 2. MZP	232	172	37	23	0.000
ELG 3. MZP - ELG 1. MZP	193	107	44	42	0.000
ELG 4. MZP - ELG 2. MZP	232	132	51	49	0.000
ELO 3. MZP - ELO 1. MZP	193	113	41	39	0.000
ELO 4. MZP - ELO 2. MZP	232	130	65	37	0.000
ERG 3. MZP - ERG 1. MZP	193	151	13	29	0.000
ERG 4. MZP - ERG 2. MZP	232	151	28	53	0.000

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, N = Anzahl, MZP = Messzeitpunkt, p = Irrtumswahrscheinlichkeit (asymptotische Signifikanz, 2-seitig)

Häufigkeitsverteilungen von Mittelwertdifferenzen machen den Lernverlauf zwischen den Messzeitpunkten sichtbar. Die Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A zwischen dem 1. und 3. sowie zwischen dem 2. und 4. Messzeitpunkt sind in Abbildung Y1 im Anhang Y dargestellt. Zusammenfassend gibt Tabelle 63 die Mittelwerte und Standardabweichungen sowie das Minimum und Maximum für die Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A wieder.

Tabelle 63: Kennwerte der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A

	N	M	SD	Min.	Max.
OG 3. MZP - OG 1. MZP	193	3.8	4.1	-9	15
OG 4. MZP - OG 2. MZP	232	3.4	3.9	-10	13
ELG 3. MZP - ELG 1. MZP	193	1.0	2.3	-6	9
ELG 4. MZP - ELG 2. MZP	232	1.0	2.3	-4	13
ELO 3. MZP - ELO 1. MZP	193	1.3	2.6	-8	10
ELO 4. MZP - ELO 2. MZP	232	1.0	2.6	-6	15
ERG 3. MZP - ERG 1. MZP	193	1.7	1.5	-2	5
ERG 4. MZP - ERG 2. MZP	232	1.2	1.6	-3	6

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, N = Anzahl, M = Mittelwert, SD = Standardabweichung, Min. = Minimum, Max. = Maximum

Wie in Tabelle 63 zu erkennen, sind die gemittelten Leistungsveränderungen sehr niedrig. Die Veränderungen von Messzeitpunkt zu Messzeitpunkt mit 1 bis 4 mehr richtig geschriebenen Wörtern und Lupenstellen sind, bis auf den orthographischen Gesamtwert, nur unwesentlich. Eine deutliche Lernentwicklung und Leistungsveränderung ist somit nur für den orthographischen Gesamtwert erkennbar.

Zusammenfassend betrachtet, zeigt sich, dass der orthographische Gesamtwert der Form A dazu geeignet ist, den Lernverlauf zu erfassen. Die Subskalen hingegen sind dazu nicht geeignet. Einerseits gibt es zwischen den Messzeitpunkten der Subskalen signifikante Unterschiede, die Leistungsveränderungen deutlich machen. Andererseits sind die absoluten Häufigkeiten der Mittelwertdifferenzen so niedrig, dass keine Schätzung des Lernverlaufs möglich ist. Das bedeutet, dass die Form A Lernentwicklungen und Leistungsveränderungen nur mit dem orthographischen Gesamtwert in ausreichendem Maße erfassen kann.

5.2.5.2 Häufigkeitsveränderungen und Mittelwertdifferenzen der Form B

Die Häufigkeitsverteilungen für den orthographischen Gesamtwert und die Subskalen der Form B zwischen dem 1. und 3. sowie zwischen dem 2. und 4. Messzeitpunkt sind in den Abbildungen 30, 31, 32 und 33 dargestellt.

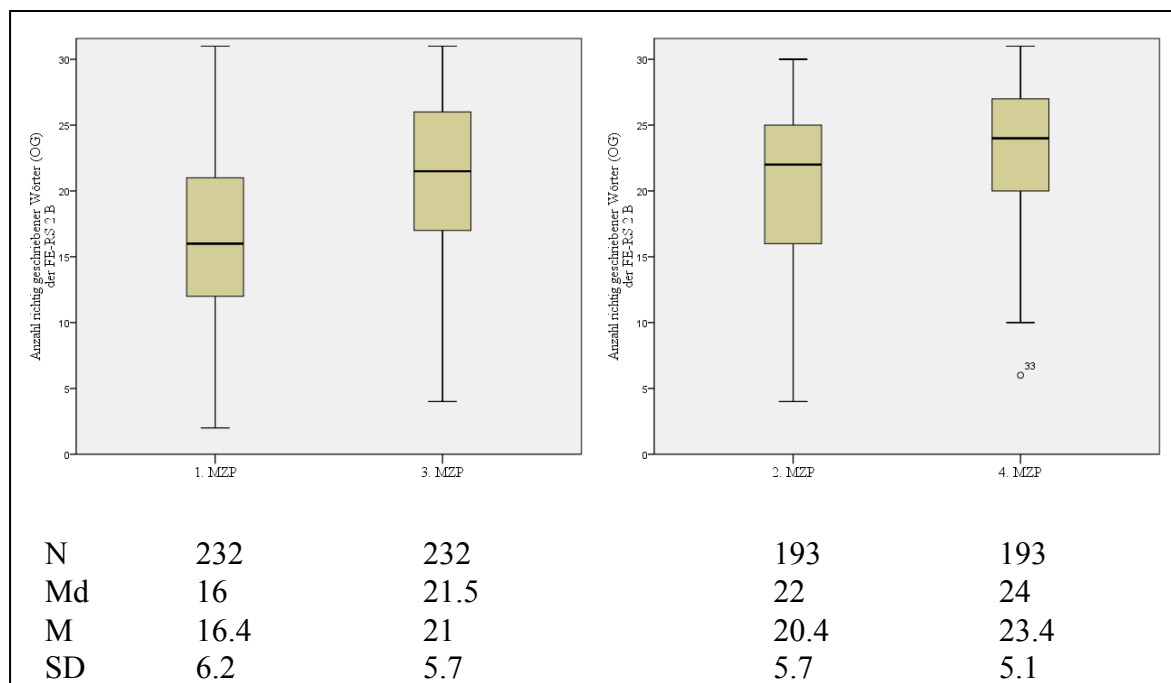


Abbildung 30: Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

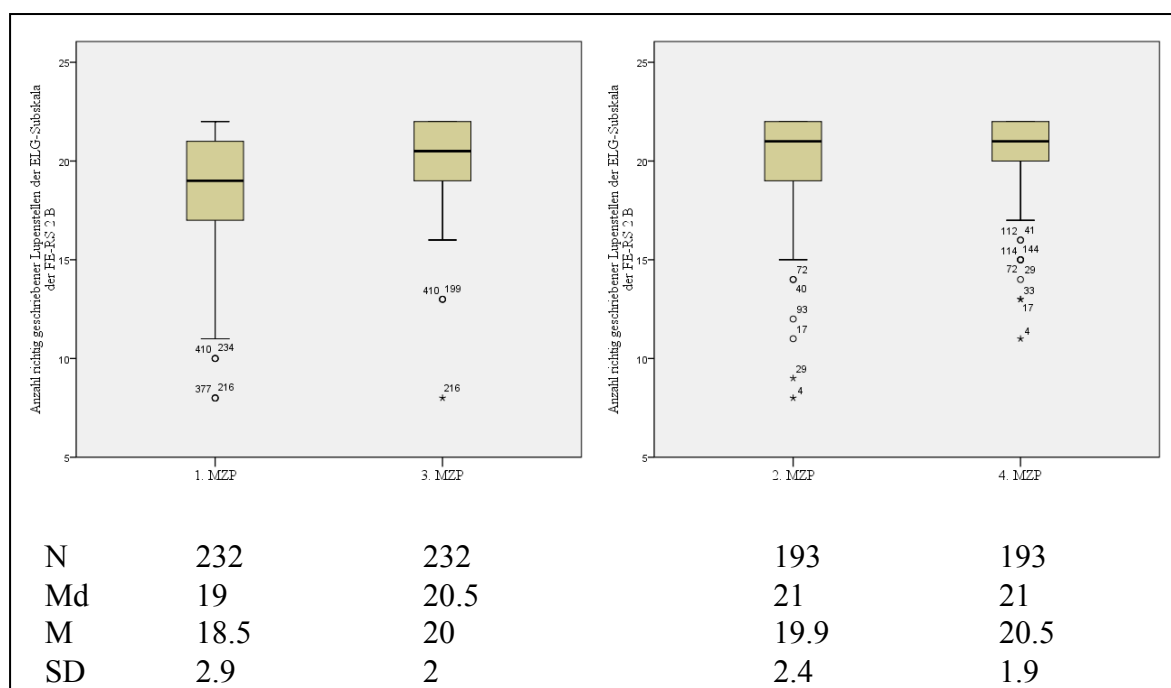


Abbildung 31: Häufigkeitsverteilungen für die ELG-Subskala der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELG = elementar graphematisches Können, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

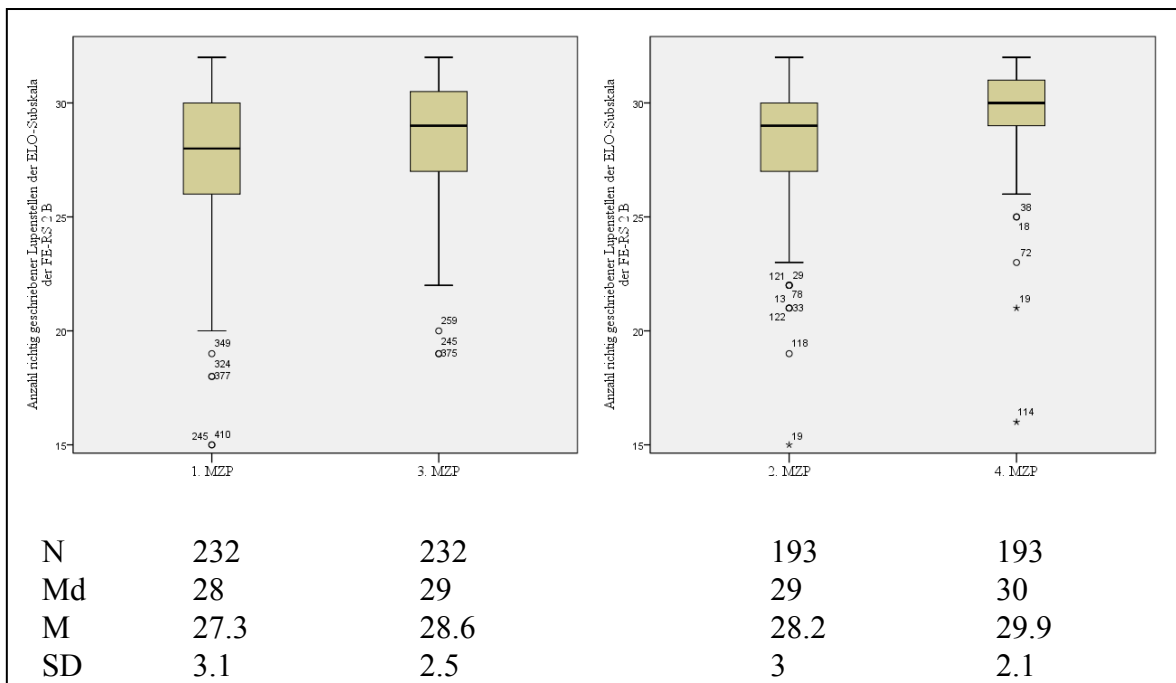


Abbildung 32: Häufigkeitsverteilungen für die ELO-Subskala der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELO = elementar orthographisches Können, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

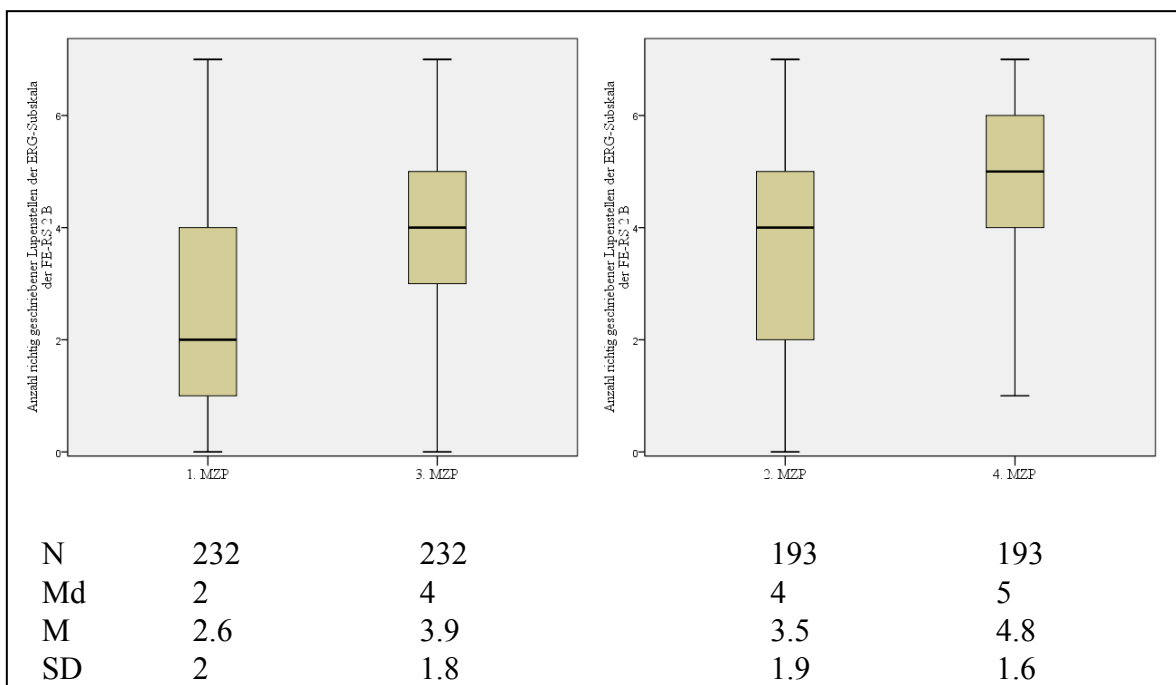


Abbildung 33: Häufigkeitsverteilungen für die ERG-Subskala der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ERG = erweitert graphematisches Können, N = Anzahl, Md= Median, M = Mittelwert, SD = Standardabweichung

Wie in den Abbildungen 30 bis 33 zu erkennen, überschneiden sich die

Interquartilbereiche der Form B, vor allem der ELG- und ELO-Subskalen, sehr stark. Auch die Mediane und Mittelwerte unterscheiden sich, bis auf den orthographischen Gesamtwert, kaum zwischen den Messzeitpunkten. Somit sind augenscheinlich deutliche Veränderungen nur bei dem orthographischen Gesamtwert erkennbar. Die Standardabweichungen des orthographischen Gesamtwerts und der Subskalen verkleinern sich von Messzeitpunkt zu Messzeitpunkt. Die Merkmalsausprägungen streuen somit von Messzeitpunkt zu Messzeitpunkt weniger stark und die Stichprobe wird hinsichtlich der Rechtschreibleistungen etwas homogener.

Der Vergleich der Mittelwerte der Form B wurde mit dem Wilcoxon-Test durchgeführt, da keine Normalverteilungen vorlagen. Die Ergebnisse des Mittelwertvergleichs zeigen höchst signifikante Unterschiede zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt des orthographischen Gesamtwerts und der Subskalen. Vom 1. zum 3. sowie vom 2. zum 4. Messzeitpunkt steigen für den orthographischen Gesamtwert und die Subskalen der Form B die Anzahlen richtig geschriebener Wörter und Lupenstellen in mindestens 46 % der Fälle. Damit ist in diesen Fällen ein Lernfortschritt im Rechtschreiben zu erkennen. In maximal 20 % der Fälle sinkt und in maximal 34 % der Fälle stagniert die Anzahl richtig geschriebener Wörter und Lupenstellen. In diesen Fällen ist kein Lernfortschritt im Rechtschreiben zu erkennen. Alle einzelnen positiven, negativen und Null differenzen sind in Tabelle 64 zusammengefasst.

Tabelle 64: Mittelwertvergleich des orthographischen Gesamtwerts und der Subskalen der Form B

	N	Positive Differenzen	Negative Differenzen	Null- differenzen	p
OG 3. MZP - OG 1. MZP	232	193	22	17	0.000
OG 4. MZP - OG 2. MZP	193	147	30	16	0.000
ELG 3. MZP - ELG 1. MZP	232	155	25	52	0.000
ELG 4. MZP - ELG 2. MZP	193	88	39	66	0.000
ELO 3. MZP - ELO 1. MZP	232	135	47	50	0.000
ELO 4. MZP - ELO 2. MZP	193	128	26	39	0.000
ERG 3. MZP - ERG 1. MZP	232	149	26	57	0.000
ERG 4. MZP - ERG 2. MZP	193	140	22	31	0.000

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, N = Anzahl, MZP = Messzeitpunkt, p = Irrtumswahrscheinlichkeit (asymptotische Signifikanz, 2-seitig)

Häufigkeitsverteilungen von Mittelwertdifferenzen machen den Lernverlauf zwischen den

Messzeitpunkten sichtbar. Die Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form B zwischen dem 1. und 3. sowie zwischen dem 2. und 4. Messzeitpunkt sind in Abbildung Z1 im Anhang Z dargestellt. Zusammenfassend gibt Tabelle 65 die Mittelwerte und Standardabweichungen sowie das Minimum und Maximum für die Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form B wieder.

Tabelle 65: Kennwerte der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form B

	N	M	SD	Min.	Max.
OG 3. MZP - OG 1. MZP	232	4.6	4.5	-9	16
OG 4. MZP - OG 2. MZP	193	3.0	4.0	-8	14
ELG 3. MZP - ELG 1. MZP	232	1.6	2.2	-4	11
ELG 4. MZP - ELG 2. MZP	193	0.7	1.9	-5	9
ELO 3. MZP - ELO 1. MZP	232	1.3	2.7	-9	11
ELO 4. MZP - ELO 2. MZP	193	1.7	2.7	-8	10
ERG 3. MZP - ERG 1. MZP	232	1.3	1.8	-4	6
ERG 4. MZP - ERG 2. MZP	193	1.3	1.6	-3	5

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, N = Anzahl, M = Mittelwert, SD = Standardabweichung, Min. = Minimum, Max. = Maximum

Wie in Tabelle 65 zu erkennen, sind die gemittelten Leistungsveränderungen sehr niedrig. Die Veränderungen von Messzeitpunkt zu Messzeitpunkt mit 1 bis 5 mehr richtig geschriebenen Wörtern und Lupenstellen sind, bis auf den orthographischen Gesamtwert, nur unwesentlich. Eine deutliche Lernentwicklung und Leistungsveränderung ist somit nur für den orthographischen Gesamtwert erkennbar.

Zusammenfassend betrachtet, zeigt sich, dass auch der orthographische Gesamtwert der Form B dazu geeignet ist, den Lernverlauf zu erfassen. Die Subskalen hingegen sind wie bei der Form A dazu nicht geeignet. Einerseits gibt es zwischen den Messzeitpunkten der Subskalen signifikante Unterschiede, die Leistungsveränderungen deutlich machen. Andererseits sind die absoluten Häufigkeiten der Mittelwertdifferenzen so niedrig, dass keine Schätzung des Lernverlaufs möglich ist. Das bedeutet, dass die Form B Lernentwicklungen und Leistungsveränderungen nur mit dem orthographischen Gesamtwert in ausreichendem Maße erfassen kann.

H_{6.1}: Die Mittelwerte des orthographischen Gesamtwerts und der Subskalen der Form A und der Form B steigen signifikant vom 1. zum 3. Messzeitpunkt und vom 2. zum 4. Messzeitpunkt.

Die Mittelwerte des orthographischen Gesamtwerts beider Formen der FE-RS 2 unterscheiden sich signifikant zwischen den Messzeitpunkten und die absoluten Häufigkeiten der Mittelwertdifferenzen erlauben eine Schätzung des Lernverlaufs. Die Mittelwerte der Subskalen beider Formen der FE-RS 2 unterscheiden sich auch signifikant zwischen den Messzeitpunkten, aber die absoluten Häufigkeiten der Mittelwertdifferenzen erlauben keine Schätzung des Lernverlaufs. Hypothese H_{6.1} lässt sich nur teilweise bestätigen. Folglich ist nur der orthographische Gesamtwert der FE-RS 2 dazu geeignet, Veränderungen der Rechtschreibleistungen in der 2. Klassenstufe zu erfassen.

5.3 Darstellung der Ergebnisse der Lehrerbefragung

Im Folgenden werden die Ergebnisse des Lehrerfragebogens zur Einschätzung der FE-RS 2 entsprechend der einzelnen Fragestellungen dargestellt.

Der Lehrerfragebogen zur Einschätzung der FE-RS 2 wurde zum 4. Messzeitpunkt an alle 29 teilnehmenden Deutschlehrkräfte ausgeteilt. Es wurden 23 Fragebögen (79 %) ausgefüllt.

Die Mehrheit der Lehrkräfte nutzt im Lernbereich Rechtschreiben parallel offene Ansätze, wie z. B. den Ansatz „Lesen durch Schreiben“ (Reichen, 2007) und die lehrgangsorientierte Arbeit mit Sprachbüchern. Keine Lehrkraft arbeitet ausschließlich nach einem offenen Ansatz. Tabelle 66 gibt die Häufigkeiten der Unterrichtsmethoden im Lernbereich Rechtschreiben wieder.

Tabelle 66: Häufigkeiten der Antworten zur Frage „Nach welcher Methode unterrichten Sie im Lernbereich Rechtschreiben?“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)

Offener Ansatz	0 (0)
Lehrgangsorientierte Arbeit mit Sprachbüchern	7 (30.4)
Parallele Nutzung offener Ansätze und lehrgangsorientierter Arbeit	16 (69.6)

Am häufigsten dominiert der lehrerzentrierte Unterricht den Lernbereich Rechtschreiben. Die meisten Lehrkräfte nutzen gelegentlich aber auch den offenen Unterricht. Tabelle 67 gibt die Häufigkeiten der Sozialformen im Lernbereich Rechtschreiben wieder.

Tabelle 67: Häufigkeiten der Antworten zur Frage „Welche Sozialformen verwenden Sie im Lernbereich Rechtschreiben?“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)

	häufig	gelegentlich	selten	nie
Lehrerzentrierter Unterricht	22 (95.6)	1 (4.3)	0 (0)	0 (0)
Offener Unterricht (z. B. Stations- und Freiarbeit, Tagesplan)	4 (17.4)	18 (78.3)	1 (4.3)	0 (0)

Den positiven Aussagen zur Praktikabilität und zum pädagogischen Nutzen der FE-RS 2 wurde überwiegend zugestimmt. Damit beurteilen die Lehrkräfte das Verfahren in seinem Aufbau und seiner Durchführung als positiv. Zudem ist das Verfahren bei den Lehrkräften erwünscht. Tabelle 68 gibt die Häufigkeiten der Antworten zur Praktikabilität und zum pädagogischen Nutzen der FE-RS 2 wieder.

Tabelle 68: Häufigkeiten der Antworten zur Frage „Wie bewerten Sie die FE-RS 2? Bitte äußern Sie sich zu folgenden Aussagen!“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)

	Stimmt genau	Stimmt etwas	Stimmt eher nicht	Stimmt nicht	k. A.
Der Zeitaufwand des Verfahrens ist angemessen.	9 (39.1)	8 (34.8)	3 (13)	2 (8.7)	1 (4.3)
Die Art der Durchführung ist praktikabel.	12 (52.2)	10 (43.5)	0 (0)	0 (0)	1 (4.3)
Das Wortmaterial des Verfahrens ist angemessen.	13 (56.5)	8 (34.8)	1 (4.3)	0 (0)	1 (4.3)
Der Nutzen des Verfahrens für Ihre pädagogische Arbeit ist groß.	5 (21.7)	9 (39.1)	5 (21.7)	0 (0)	4 (17.4)
Sie haben Interesse an einer Anwendung des Verfahrens im Unterricht.	3 (13)	13 (56.5)	4 (17.4)	1 (4.3)	2 (9)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, k. A. = keine Angaben

Nach Aussage der Lehrkräfte erlebten die Schüler die FE-RS 2 überwiegend als positiv. Sie hatten keine Angst vor der Durchführung der FE-RS 2 und waren auch nach der Durchführung nicht enttäuscht. Zum Teil waren die Schüler aber nach der Durchführung der FE-RS 2 erschöpft. Tabelle 69 gibt die Häufigkeiten der Antworten zum Schülerverhalten wieder.

Tabelle 69: Häufigkeiten der Antworten zur Frage „Wie bewerten die Schüler Ihrer Meinung nach die FE-RS 2? Bitte äußern Sie sich zu folgenden Aussagen!“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)

	Stimmt genau	Stimmt etwas	Stimmt eher nicht	Stimmt nicht	k. A.
Die Schüler erleben das Verfahren positiv.	5 (21.7)	18 (78.3)	0 (0)	0 (0)	0 (0)
Die Schüler weigerten oder ängstigten sich, das Verfahren durchzuführen.	0 (0)	2 (8.7)	2 (8.7)	18 (78.3)	1 (4.3)
Die Schüler waren nach der Durchführung des Verfahrens erschöpft.	3 (13)	9 (39.1)	10 (43.5)	0 (0)	1 (4.3)
Die Schüler waren nach der Durchführung des Verfahrens enttäuscht oder demotiviert.	0 (0)	1 (4.3)	13 (56.5)	8 (34.8)	1 (4.3)

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, k. A. = keine Angaben

Zusammenfassend betrachtet, ist die FE-RS 2 ein Verfahren zur Erfassung der Rechtschreibleistung im 2. Schuljahr, das Lehrkräfte positiv bewerten und an dem sie Interesse an der Anwendung im Unterricht haben.

6 Diskussion

Die vorliegende Arbeit sollte einen Beitrag zur frühen und entwicklungsbezogenen Diagnostik von Rechtschreibleistungen leisten und die Frage beantworten, ob die FE-RS 2 ein valides, reliables und veränderungssensitives Verfahren ist, das die Lernentwicklung und den Lernverlauf der Schüler im Lernbereich Rechtschreiben der 2. Klassenstufe erfassen kann. In diesem Kapitel werden die umfangreichen Ergebnisse der Untersuchung zusammengefasst und diskutiert. Anschließend werden die Grenzen der Untersuchung dargelegt und es folgt ein Vergleich der FE-RS 2 mit bestehenden Verfahren der Rechtschreibdiagnostik. In einem Ausblick wird diskutiert, wie sich die Ergebnisse der Untersuchung für die weitere Forschung und Praxis nutzen lassen.

6.1 Zusammenfassung und Diskussion der Ergebnisse

Im folgenden Abschnitt wird die testtheoretische Güte der FE-RS 2 diskutiert. Es wird diskutiert, ob die beiden Formen A und B sowie die Subskalen der FE-RS 2 valide, reliabel und veränderungssensitiv sind. In diesem Zusammenhang erfolgt ein Bezug zu den theoretischen Annahmen, die zur Entwicklung der FE-RS 2 geführt haben, um die Konstruktion des Verfahrens und die Itemauswahl kritisch zu reflektieren. Die Zusammenfassung und Diskussion der Ergebnisse erfolgt entsprechend der Reihenfolge der Forschungsfragen. Der Abschnitt schließt mit einer Anwendungsempfehlung für die FE-RS 2.

Die erste Forschungsfrage umfasste die Differenzierungsfähigkeit der FE-RS 2. Zur Schätzung der Differenzierungsfähigkeit wurden die Häufigkeitsverteilungen betrachtet sowie die Schwierigkeiten und Trennschärfen des orthographischen Gesamtwerts und der Subskalen berechnet. Die Betrachtung der Häufigkeitsverteilungen ergab, dass die Verteilungen für den orthographischen Gesamtwert der Form A, bis auf den 3. Messzeitpunkt, einer Normalverteilung entsprechen. Die Verteilungen für den orthographischen Gesamtwert der Form B entsprechen alle, bis auf den 1. Messzeitpunkt, nicht einer Normalverteilung, da sie linksschief sind. Die Verteilungen für die Subskalen der Form A über alle Messzeitpunkte entsprechen nicht einer Normalverteilung. Sie sind alle, bis auf die ERG-Subskala zum 1. und 2. Messzeitpunkt, linksschief verteilt. Die

Verteilungen für die Subskalen der Form B über alle Messzeitpunkte entsprechen ebenso nicht einer Normalverteilung. Auch sie sind alle, bis auf die ERG-Subskala zum 1., 2. und 3. Messzeitpunkt, linksschief verteilt. Zusammenfassend betrachtet, zeigten sich über alle 4 Messzeitpunkte linksschiefe Verteilungen bzw. Verschiebungen von Normalverteilungen zu linksschiefen Verteilungen. Die Schüler schrieben von Messzeitpunkt zu Messzeitpunkt mehr Wörter und Lupenstellen orthographisch richtig. Dieses Ergebnis entspricht den Annahmen des Entwicklungsmodells des graphematischen und orthographischen Rechtschreibkönnens, das die theoretische Grundlage der FE-RS 2 bildet. Aufgrund der relativ hohen Standardabweichungen im Vergleich zu den Standardmessfehlern des orthographischen Gesamtwerts der FE-RS 2 der Form A und B über alle Messzeitpunkte kann geschlossen werden, dass der orthographische Gesamtwert der FE-RS 2 zwischen Schülern mit unterschiedlichen Rechtschreibleistungen differenzieren kann. Die Subskalen der FE-RS 2 beider Formen können nicht in ausreichendem Maße zwischen unterschiedlich leistungsstarken Schülern differenzieren. Sie sind aufgrund ihrer linksschiefen Verteilungen eher dazu geeignet, im unteren Leistungsbereich zu differenzieren und Schüler mit schwachen Rechtschreibleistungen zu erkennen. Dieses Ergebnis entspricht der Intention der FE-RS 2, Schüler mit Schwierigkeiten im Rechtschreiberwerb so früh wie möglich zu identifizieren und deren Rechtschreibschwierigkeiten qualitativ einschätzen zu können. Die Analyse der Schwierigkeiten ergab, dass die FE-RS 2 beider Formen unterschiedliche Schwierigkeitsgrade abdeckt und die Schwierigkeitsindexe für den orthographischen Gesamtwert und die Subskalen innerhalb der FE-RS 2 der Form A und B steigen. Die Schwierigkeitsindexe für den orthographischen Gesamtwert und die Subskalen der FE-RS 2 beider Formen steigen über alle 4 Messzeitpunkte. Die Schüler schreiben also von Messzeitpunkt zu Messzeitpunkt mehr Wörter und Lupenstellen richtig. Diese Ergebnisse entsprechen der Konzeption der FE-RS 2, nämlich, dass die Wörter bezüglich der Rechtschreibbesonderheiten innerhalb des Verfahrens schwieriger werden, von Messzeitpunkt zu Messzeitpunkt aber leichter zu schreiben sind. Aufgrund der mittleren Itemschwierigkeiten können der orthographische Gesamtwert und die ERG-Subskala beider Formen der FE-RS 2 relativ gut differenzieren. Die ELG- und ELO-Subskalen sind wegen des hohen Schwierigkeitsindexes über alle 4 Messzeitpunkte für eine gute Differenzierung im unteren Leistungsbereich geeignet, was der Intention der FE-RS 2 entspricht. Die Analysen der Trennschärfen entsprechen den Ergebnissen der Itemschwierigkeiten. Insgesamt stellt sich die Differenzierungsfähigkeit der FE-RS 2

unterschiedlich dar. Der orthographische Gesamtwert ist am besten dazu geeignet, Leistungsunterschiede abzubilden, wobei die Form A dabei der Form B aufgrund der Normalverteilung überlegen ist. Die Subskalen beider Formen der FE-RS 2 sind vorwiegend dazu geeignet, im unteren Leistungsbereich zu differenzieren.

Die zweite Forschungsfrage umfasste die Reliabilität der FE-RS 2. Zur Schätzung der Reliabilität wurden die Retest- und Testhalbierungsreliabilitäten und die internen Konsistenzen berechnet. Die Berechnungen ergaben, dass die Reliabilitätswerte für die Retestreliabilität der FE-RS 2 beider Formen für den orthographischen Gesamtwert im hohen und für die Subskalen überwiegend im mittleren Bereich liegen. Für die Korrelationen der Testhalbierungsreliabilität und die internen Konsistenzen ergaben sich ähnliche Ergebnisse. Für beide Formen der FE-RS 2 liegt der orthographische Gesamtwert im hohen und für die Subskalen im mittleren bis hohen Bereich. Insgesamt sind damit die Reliabilitäten des orthographischen Gesamtwerts und der Subskalen für die Formen A und B zufriedenstellend. Sowohl der orthographische Gesamtwert als auch die Subskalen der FE-RS 2 messen genau. Im Vergleich zu Studien zur Reliabilität von Rechtschreib-CBMs sind die vorliegenden Daten mit hohen Reliabilitäten für den orthographischen Gesamtwert und mit überwiegend mittelhohen Reliabilitäten für die Subskalen der FE-RS 2 jedoch zum Teil unbefriedigend. Die von Marston (1989) zitierten Studien von Marston (1982), Tindal, Germann und Deno (1983) sowie M. R. Shinn (1981) berichten über Retestreliabilitäten zwischen $r_{tt} = .85$ und $r_{tt} = .94$. In der vorliegenden Arbeit findet sich nur ein ähnlich hoher Retestwert mit $r_{tt} = .79$ für den orthographischen Gesamtwert der Form A zwischen dem 2. und 4. Messzeitpunkt. Trotz eines ähnlich langen Retestintervalls kann die HSP (May, 2002) zwischen der Mitte (HSP 1+ M2) (May, 2005a) und dem Ende der 2. Klassenstufe (HSP 2) (May, 2005b) aber, wie die oben genannten Studien, für die Graphemtreffer eine sehr hohe Retestreliabilität von $r_{tt} = .93$ aufweisen. Wodurch diese großen Unterschiede zwischen den Retestreliabilitäten verursacht sind und warum die FE-RS 2 insgesamt nur mit mittlerer Reliabilität misst, kann nur vermutet werden. Wahrscheinlich ist es aber so, dass die Itemanzahl die Reliabilität stark beeinflusst (Bühner, 2004). Die FE-RS 2 wurde als Screening entwickelt, das einen Überblick über die Rechtschreibleistung der Schüler geben und zugleich zeitökonomisch sein sollte. Dieser Anspruch ging zulasten der Itemanzahl, die FE-RS 2 besteht aus nur 31 Wörtern. Zwar besteht die HSP 1+ M2 (May, 2005a) auch nur aus 15 Wörtern und die HSP 2 (May, 2005b) aus 30 Wörtern, aber die Grundlage für die Auswertung der HSP (May, 2002) und

die Berechnungen der Retestreliabilitäten sind die Graphemtreffer. Da die HSP 1+ M2 (May, 2005a) 63 und die HSP 2 (May, 2005b) 148 Graphemtreffer hat, wirkt sich diese hohe Anzahl vermutlich sehr positiv auf die Höhe der Reliabilitäten aus. Dass die FE-RS 2 zugunsten einer besseren Reliabilität um Items ergänzt werden sollte, entspricht nicht dem zeitökonomischen Anspruch des Verfahrens und wird nicht in Erwägung gezogen. Grund dafür ist vor allem die Sorge, das zeitintensive Verfahren vermutlich kaum von Lehrkräften genutzt werden würden. Insgesamt bleibt festzuhalten, dass die FE-RS 2, unabhängig von der Form A oder B, dazu geeignet ist, zuverlässige und präzise Aussagen über den Leistungsstand der Schüler zu treffen.

Die dritte Forschungsfrage umfasste die Validität der FE-RS 2. Zur Schätzung der Validität wurden die Inhalts- und Übereinstimmungsvalidität sowie die konkurrente, diskriminante und faktorielle Validität bestimmt. Die Ergebnisse sprechen für eine zufriedenstellende Validität des Verfahrens. Zwischen dem orthographischen Gesamtwert der FE-RS 2 und der Rechtschreibnote liegt für die Form A eine hohe, für die Form B eine mittelhohe Korrelation vor. Der orthographische Gesamtwert der Form A korreliert mit dem DERET 1-2+ (Stock & Scheider, 2008a) hoch, die Form B mittel bis hoch. Die konkurrente Validität der FE-RS 2 mit der FE-L 2 (Kuhlmann & Hartke, 2009b) ist zufriedenstellend. Die sehr geringen Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2 und der FE-AF 2 (Kuhlmann & Hartke, 2009a), des DEMAT 1+ (Krajewski et al., 2002) sowie des DEMAT 2+ (Krajewski et al., 2004) bestätigen zudem, dass es sich bei dem gemessenen Merkmal der FE-RS 2 nicht um arithmetische Fähigkeiten handelt, sondern um eine andere Leistung. Insgesamt zeigen die Ergebnisse zur Validitätsprüfung, dass der orthographische Gesamtwert der FE-RS 2 Rechtschreibleistungen valide erfassen kann. Die in dieser Arbeit ermittelten Ergebnisse sind mit den Ergebnissen von Studien zur Validität von Rechtschreib-CBMs teilweise vergleichbar. Die von Marston (1989) zitierten Studien von Deno, Mirkin, Lowry und Kuehnle (1980) sowie Marston (1982) berichten über kriteriumsorientierte Validitäten zwischen $r = .80$ und $r = .99$. Auch in der vorliegenden Arbeit findet sich mit einer Korrelation von $r = -.81$ zwischen der Form A und dem DERET 1-2+ (Stock & Schneider, 2008a) zum 4. Messzeitpunkt ein ähnlich hoher Validitätswert. Die statistische Prüfung auf Brauchbarkeit der postulierten Subskalen der FE-RS 2 erfolgte mittels Faktorenanalyse. Die Subskalen der FE-RS 2 wurden aufgrund theoretischer Überlegungen zum Entwicklungsmodell des Rechtschreibkönnens, das dieser Arbeit zugrunde liegt, gebildet.

Die FE-RS 2 wurde so konstruiert, dass vorwiegend elementar graphematisches und orthographisches Können erfasst werden sollte. Zudem wurden zwei Skalen für erweitert graphematisches und orthographisches Können gebildet. Das Ergebnis der Faktorenanalyse zeigte aber, dass die Items nicht nach den vier Subskalen gruppiert wurden. Die Faktorenanalyse für die Form A ergab eine relativ eindeutige Zuordnung der Items zu zwei Faktoren. Diese zwei Faktoren, die als elementare und erweiterte Stufen bezeichnet werden können, entsprechen der Grundstruktur des Entwicklungsmodells graphematischen und orthographischen Rechtschreibkönnens, welches die theoretische Grundlage für die Konzeption der FE-RS 2 ist. Für die Form B ist aufgrund der Ergebnisse der Faktorenanalyse diese Grundstruktur, bis auf die elementare Stufe, nicht erkennbar. In Bezug zur faktoriellen Validität kann geschlussfolgert werden, dass die inhaltlich begründeten Subskalen der FE-RS 2 für die Form A in ihrer Grundstruktur statistisch vertretbar sind. Zwar sind nicht alle vier Subskalen erkennbar, jedoch die zwei Ausprägungsebenen des Rechtschreibkönnens, die elementare und erweiterte Stufe, die ursprünglich dem Kompetenzentwicklungsmodell des Rechtschreibens von Löffler und Meyer-Schepers (2008) entstammen. Dass die Ergebnisse der Faktorenanalyse nicht die vier theoretisch vertretbaren Subskalen stützen, liegt vermutlich an der Bildung von „*Schwierigkeitsfaktoren*“ (Lienert & Raatz, 1998, S. 113). Neben der Gruppierung von Items nach inhaltlichen Ähnlichkeiten können Items bei der Faktorenanalyse nämlich auch nach ähnlichen Schwierigkeiten zu Faktoren gruppiert werden. Diese sogenannten Schwierigkeitsfaktoren erschweren dann die Interpretation der Faktorenstruktur (Amelang, Bartussek, Stemmler & Hagemann, 2006; Lienert & Raatz, 1998). Trotz der zum Teil unbefriedigenden Ergebnisse der Faktorenanalyse sollten die Subskalen der FE-RS 2 bestehen bleiben, da sie Hinweise darauf geben, welche Rechtschreibstrategien Schüler beherrschen und welche nicht. Insgesamt ist die Validität der FE-RS 2 für beide Formen zufriedenstellend. Aufgrund der etwas höheren Korrelationen bezüglich der Konstruktvalidität und der statistisch vertretbaren Grundstruktur der Subskalen ist die Form A der Form B ein wenig überlegen und besser dazu geeignet, valide zu messen.

Die vierte Forschungsfrage umfasste die prognostische Validität der FE-RS 2. Zur Schätzung der prognostischen Aussagekraft wurden Regressionsanalysen und klassifikatorische Analysen durchgeführt. Die Regressionsanalysen und klassifikatorischen Analysen ergaben, dass die Vorhersageleistungen des orthographischen Gesamtwerts der Formen A und B zum 1. Messzeitpunkt für die Rechtschreibleistung zum 3. Messzeitpunkt

annähernd gleich hoch sind. Die Vorhersageleistungen des orthographischen Gesamtwerts beider Formen zum 2. Messzeitpunkt für die Rechtschreibleistungen zum 4. Messzeitpunkt unterscheiden sich voneinander. Die Regressionsanalysen ergaben, dass der orthographische Gesamtwert der Form A zum 2. Messzeitpunkt mit 62 % Varianzaufklärung am besten dazu geeignet ist, die Rechtschreibleistung am Ende der 2. Klassenstufe vorherzusagen. Die Varianzaufklärung der Form B ist niedriger. Die klassifikatorischen Analysen bestätigen mit einem RAZ-Index von 72 % für den orthographischen Gesamtwert der Form A zum 2. Messzeitpunkt das Ergebnis der Regressionsanalyse. Somit können mit dem orthographischen Gesamtwert der Form A zum Ende des ersten Schulhalbjahres der 2. Klassenstufe 41 % der Schüler mit zukünftigen Rechtschreibschwierigkeiten richtig vorausgesagt werden. Die Vorhersageleistung der 16 % der Schwächsten im DERET 1-2+ (Stock & Schneider, 2008a) am Ende der 2. Klassenstufe ist mit dem orthographischen Gesamtwert der Form A mit RAZ-Indizes von 69 % zum 1. Messzeitpunkt und von 90 % zum 3. Messzeitpunkt ebenfalls besser als für die Form B. Die Subskalen der FE-RS 2 sind, bis auf die ELO-Subskala der Form B zum 2. Messzeitpunkt, zu keinem Messzeitpunkt für eine sehr gute Vorhersage des orthographischen Gesamtwerts zum 4. Messzeitpunkt geeignet. Ebenso können der orthographische Gesamtwert und die Subskalen beider Formen der FE-RS 2 nicht die Rechtschreibnote am Schuljahresende vorhersagen. Dass die Subskalen der FE-RS 2 nicht für genaue Vorhersagen der Rechtschreibleistungen geeignet sind, lässt sich mit der unterschiedlichen inhaltlichen Konzeption der Skalen begründen. Die Skala „Orthographischer Gesamtwert“ umfasst Wörter unterschiedlicher Schwierigkeitsgrade, die das gesamte Konstrukt Rechtschreibleistung erfassen. Die Subskalen hingegen umfassen Lupenstellen, die einzelne orthographische Besonderheiten in Wörtern markieren und nur einen kleinen Teilbereich des Gesamtkonstrukts Rechtschreibleistung erfassen. Die Lupenstellen der Subskalen bilden demnach nur eine kleine Teilmenge der Gesamtmenge „Orthographischer Gesamtwert“. Die Schnittmenge zwischen einer Subskala und dem orthographischen Gesamtwert ist so klein, dass eine Subskala keine genauen Vorhersagen zum orthographischen Gesamtwert zulässt. Dass die FE-RS 2, trotz mittlerer und hoher Korrelationen zwischen der Rechtschreibnote und dem orthographischen Gesamtwert, nicht für eine genaue Vorhersage der Rechtschreibnote geeignet ist, liegt eventuell an fehlerhaften oder verzerrten Noten (Schabmann & B. M. Schmidt, 2009). Insgesamt bleibt festzuhalten, dass der orthographische Gesamtwert der Form A ab dem Ende des ersten Schulhalbjahres der 2. Klassenstufe genaue Vorhersagen

zu den Rechtschreibleistungen am Schuljahresende treffen kann. Der Wunsch einer möglichst frühen Diagnostik von Rechtschreibleistungen und der damit verbundene Anspruch, bereits zu Beginn der 2. Klassenstufe oder noch eher valide Vorhersagen zu den erwarteten Rechtschreibleistungen treffen zu können, ist mit der FE-RS 2 jedoch nicht valide realisierbar und außerdem zu hinterfragen. Es stellt sich nämlich die Frage, ab welchem Zeitpunkt rechtschriftliche Fähigkeiten erwartet werden können. Nach Mannhaupt (2008), Scheerer-Neumann (2007, 2008) und dem dieser Arbeit zugrunde liegenden Entwicklungsmodell graphematischen und orthographischen Rechtschreibkönnens gewinnen Schüler in der 2. Klasse, nach der alphabetischen Phase, erste Einsichten in orthographische Regelmäßigkeiten und Strukturen. Sodann könnten mit Beginn des Rechtschreibunterrichts in der 2. Klassenstufe erste rechtschriftliche Fähigkeiten erworben und gefestigt werden. Dieser beginnende Aneignungsprozess des Rechtschreibens benötigt viel Zeit zum Einführen, Wiederholen und Üben, sodass vermutlich davon ausgegangen werden kann, erste rechtschriftliche Fähigkeiten zum Ende des ersten Schulhalbjahres der 2. Klassenstufe abprüfen zu können. Insofern wäre der erste Einsatz der FE-RS 2 in der 10. Schulwoche zu früh. In Anbetracht der mittleren Itemschwierigkeiten des orthographischen Gesamtwerts, der niedrigen bis mittleren Itemschwierigkeiten der Subskalen und der hohen Korrelationen zwischen dem orthographischen Gesamtwert beider Formen der FE-RS 2 und dem DERET 1-2+ (Stock & Schneider, 2008a) zum 1. Messzeitpunkt scheint der gewählte erste Anwendungszeitraum jedoch nicht zu früh. Zudem gibt auch das prognostische Ergebnis der FE-RS 2 zum 1. Messzeitpunkt einen ersten Hinweis auf die Notwendigkeit differenzierender Maßnahmen im Unterricht. Des Weiteren ist, vor dem Hintergrund der Einschätzung des Schülerverhaltens während der Durchführung der FE-RS 2 durch die Lehrkräfte, davon auszugehen, dass sich ein früher Einsatz der FE-RS 2 nicht nachteilig auf das Verhalten der Schüler auswirkt. Schlussfolgernd kann die Form A, vorrangig der orthographische Gesamtwert, bereits ab der 10. Schulwoche der 2. Klassenstufe als Screening eingesetzt werden.

Die fünfte Forschungsfrage umfasste die Äquivalenz der Formen A und B der FE-RS 2. Zur Schätzung der Äquivalenz wurden die Varianzen, Mittelwerte, Häufigkeitsverteilungen sowie die Reliabilitäts- und Validitätswerte beider Formen miteinander verglichen. Die Berechnungen ergaben, dass beide Formen der FE-RS 2 hinsichtlich der Validität und Reliabilität überwiegend gleich korrelieren und auch die

Werte der Paralleltestreliabilität für den orthographischen Gesamtwert hoch sind. Neben der Forderung nach gleichen Mittelwerten, Varianzen und Häufigkeitsverteilungen kann vor allem aber die primäre Forderung nach hoher Paralleltestreliabilität für die gesamte FE-RS 2, das heißt orthographischer Gesamtwert und Subskalen, nicht erfüllt werden. Aufgrund dessen scheint sich die der Konstruktion der FE-RS 2 zugrunde liegende Annahme, nämlich dass Wortpaare mit sehr ähnlichen linguistischen und rechtschreiblichen Besonderheiten, z. B. Qualm- Quark oder Wolke- Hefte, in ihrem Schwierigkeitsgrad identisch sein müssten, nicht zu bestätigen. Auch Strathmann und Klauer (2008) sowie Strathmann, Klauer und Greisbach (2010) können in ihrer Pilotstudie zur Entwicklung eines Verfahrens der Lernverlaufsdiagnostik für die Rechtschreibkompetenz nur überwiegend unbefriedigende Paralleltestreliabilitäten vorweisen. Strathmann und Klauer (2008) vermuten, dass die schwachen Reliabilitäten zum einen durch eine Fluktuation der Schüler begründet sein können. Diese Vermutung kann für die vorliegende Arbeit ausgeschlossen werden, da nur vollständige Datensätze berücksichtigt wurden. Zum anderen vermuten Strathmann und Klauer (2008), dass auch emotional-motivationale Prozesse sowie individuelle Beeinträchtigungen Einfluss auf die Reliabilität der Testdiktate haben könnten. Diese Faktoren konnten in der vorliegenden Arbeit nicht kontrolliert werden und hätten Einfluss auf die Reliabilität der FE-RS 2 haben können. Für eine reliablere Gestaltung von Verfahren der Lernverlaufsdiagnostik für die Rechtschreibkompetenz schlagen Strathmann und Klauer (2008) sowie Strathmann, Klauer und Greisbach (2010) vor, die Itemanzahl auf über 20 Wörter zu erhöhen und die Wörter aus definierten Teilmengen zu ziehen. Dieses Vorgehen wurde bereits für die Entwicklung der FE-RS 2 angewandt. Dennoch ist die Paralleltestreliabilität für die gesamte FE-RS 2 nicht hinreichend hoch. Die Folge ist, dass nur die Form A oder Form B der FE-RS 2 in einer Klasse verwendet werden sollte. Unerwünschte Wiederholungs- und Erinnerungseffekte, die Testergebnisse beeinflussen, sollten aufgrund des großen Zeitintervalls von 10 Wochen zwischen den Messungen beim Einsatz nur einer Form der FE-RS 2 nicht auftreten.

Die sechste Forschungsfrage umfasste die Veränderungssensitivität der FE-RS 2. Zur Schätzung der Veränderungssensitivität wurden die Häufigkeitsveränderungen betrachtet und die Mittelwertdifferenzen berechnet. Die Berechnungen ergaben, dass sich die Mittelwerte des orthographischen Gesamtwerts und der Subskalen der FE-RS 2 zwischen den Messzeitpunkten signifikant voneinander unterscheiden. Die absoluten Häufigkeiten

der Mittelwertdifferenzen lassen nur für den orthographischen Gesamtwert, nicht aber für die Subskalen, eine Schätzung des Lernverlaufs zu. Damit ist der orthographische Gesamtwert der FE-RS 2 in der Lage, Veränderungen der Rechtschreibleistungen in der 2. Klassenstufe zu erfassen. Aufgrund der „klassischen Probleme der Veränderungsmessung“ (Rost, 1996, S. 260) ist die Veränderungssensitivität des orthographischen Gesamtwerts der FE-RS 2 aber eingeschränkt. Die klassischen Probleme der Veränderungsmessung betreffen das Maß für Veränderungen, die Differenzwerte.

„Mit diesen Differenzwerten sind drei Probleme verbunden, nämlich

1. sie sind meistens sehr unreliabel, also mit einem hohen Messfehler behaftet,
2. sie korrelieren negativ mit der Personenvariable zum Zeitpunkt $t=1$, d. h. [das heißt] der sogenannte Anfangswert und der Differenzwert sind negativ korreliert und
3. es stellt sich die Frage, ob man diese Differenzen überhaupt bilden darf, da man bekanntlich nur Gleiches von Gleichem abziehen darf. Es ist die Frage zu beantworten, ob Vor- und Nachtest dasselbe messen“ (Rost, 1996, S. 260).

Insofern lassen sich mit der FE-RS 2 der Lernverlauf und die Veränderungen der Rechtschreibleistungen nur vorsichtig und annähernd in Bezug zur Stichprobe schätzen. Valide und reliable Lernverlaufs- und Veränderungsmessungen sind mit der FE-RS 2 nicht möglich.

Zusammenfassend betrachtet, legen die Ergebnisse der Untersuchung nahe, die FE-RS 2 in der Praxis anzuwenden. Allerdings sollte die FE-RS 2 nicht in der vorliegenden konzipierten Form angewendet, sondern aufgrund der Ergebnisse der Untersuchung verändert werden. Es sollte nur die Form A oder B der FE-RS 2 verwendet werden, da die Paralleltestreliabilität des Verfahrens nicht zufriedenstellend ist. Für die Anwendung in der Praxis wird empfohlen, die FE-RS 2 der Form A zu verwenden, da diese Form eine bessere Differenzierungsfähigkeit hat, höhere Werte für die Konstruktvalidität und prognostische Validität aufweist und die theoretisch begründete Grundstruktur des Verfahrens statistisch bestätigt wird. Sodann ist die Form A ein objektives, reliables und valides Verfahren, das zu vier Messzeitpunkten die Rechtschreibleistungen von Zweitklässlern erfasst. Der orthographische Gesamtwert des Verfahrens bildet am besten die Leistungsunterschiede im Rechtschreiben ab. Insbesondere die Subskalen differenzieren aber im unteren Leistungsbereich, sodass Schüler mit geringen Rechtschreibleistungen identifiziert werden können. Der orthographische Gesamtwert der Form A sagt bereits zum Ende des ersten Schulhalbjahres der 2. Klassenstufe die Rechtschreibleistungen valide voraus und ist durch die mehrmaligen Messwiederholungen dazu geeignet, den Lernverlauf im Rechtschreiben

in der 2. Klassenstufe vorsichtig zu schätzen. Insgesamt gibt die Form A Lehrkräften wichtige Informationen über das rechtschriftliche Können der Schüler zu mehreren Messzeitpunkten in der 2. Klassenstufe.

6.2 Fazit und Ausblick

In diesem Abschnitt wird zuerst die antizipierte Anwendung der FE-RS 2 im Kontext von Prävention im RTI-Ansatz beschrieben. Anschließend werden die methodischen Grenzen der vorliegenden Untersuchung diskutiert und es folgt ein Vergleich der FE-RS 2 mit anderen Rechtschreibtests.

Im Rahmen der vorliegenden Arbeit wurde vor dem Hintergrund schulischer Prävention im RTI-Ansatz die FE-RS 2 als ein curriculumnahes diagnostisches Verfahren der Lernverlaufsdokumentation für den Lernbereich Rechtschreiben der 2. Klassenstufe konzipiert. Das Ziel der FE-RS 2 ist es, die Rechtschreibleistungen von Zweitklässlern mehrmals im Schuljahr reliabel und valide zu erfassen. Die FE-RS 2 erfasst den Lernverlauf im Rechtschreiben und gibt Lehrkräften frühzeitig Rückmeldungen, bei welchem Schüler sich Rechtschreibschwierigkeiten entwickeln könnten. Im Sinne des RTI-Ansatzes gibt die FE-RS 2 auch Rückmeldungen über die Unterrichtsqualität, da mit der Messung des Lernverlaufs auch der Erfolg der angewandten pädagogischen Maßnahmen evaluiert wird. Der RTI-Ansatz zielt durch eine systematische Stärkung der Unterrichtsqualität darauf ab, auszuschließen, dass Unterricht ursächlich für Lernschwierigkeiten ist. Regelmäßig und zeitlich engmaschig werden Verfahren der Lernverlaufsdokumentation eingesetzt, um Risikoschüler so früh wie möglich zu identifizieren und durch differenzierende Maßnahmen im Unterricht oder durch zusätzliche Förderung zu fördern. Diese Anwendung im regulären Unterricht und Orientierung an der primären Prävention sind auch für die FE-RS 2 als Screeningverfahren wünschenswert. Verfahren der Lernverlaufsdokumentation, wie die FE-RS 2, bieten „in diesem Sinne die methodische Grundlage des diagnostischen Anliegens von RTI“ (C. Müller & Hartmann, 2009, S. 38), denn das diagnostische Anliegen des RTI-Ansatzes ist die Optimierung von Unterricht und Förderung durch den kontinuierlichen Einsatz von Verfahren der Lernverlaufsdokumentation.

Für den Einsatz der FE-RS 2 in der Praxis ist es wünschenswert, dass das Verfahren methodenunabhängig sowie unabhängig von pädagogischen und diagnostischen Ansätzen angewendet wird. Es wäre dann ein theoriebasiertes, formatives Verbesserungsinstrument von Rechtschreibunterricht und rechtschriftlichen Fördermaßnahmen. Auf der Grundlage der Diagnostik von Rechtschreibleistungen mit der FE-RS 2 könnten Lehrkräfte sodann datenbasiert, entsprechend der Fähigkeiten der Schüler, Unterricht adaptiv, das heißt individualisiert und differenziert, gestalten und dadurch das Auftreten von Schwierigkeiten im Rechtschreiberwerb verhindern.

Obwohl die vorliegende Untersuchung zufriedenstellende Ergebnisse vorweisen kann, weist sie methodische Grenzen auf. Nachfolgend werden die methodischen Grenzen der vorliegenden Untersuchung diskutiert.

Die *erste Einschränkung* betrifft die Rekrutierung der Stichprobe. Da die Daten nur in Mecklenburg-Vorpommern erhoben wurden, ist die Repräsentativität der vorliegenden Untersuchung eingeschränkt. So sind die Normen überregional nicht repräsentativ, sondern eher nur für Mecklenburg-Vorpommern, insbesondere den Schulamtsbezirk Neubrandenburg. Die Angaben zur Güte der FE-RS 2 sind trotz des Vorgehens aussagekräftig, da die Stichprobe ausreichend heterogen und umfangreich ist. Gründe für dieses Rekrutierungsvorgehen waren die sehr begrenzten Ressourcen, die zur Durchführung der Untersuchung zur Verfügung standen. Wenngleich vermutet wird, dass die vorliegende Untersuchung mit einer repräsentativen Stichprobe nicht zu substanziell anderen Ergebnissen führt, sollte sie in Verbindung mit einer Überarbeitung der FE-RS 2 mit einer repräsentativen Stichprobe erneut durchgeführt werden, um die Generalisierbarkeit der Ergebnisse zu prüfen und zu sichern.

Die *zweite Einschränkung* betrifft die Untersuchungsdurchführung. Aufgrund fehlender personeller Unterstützung bei der Erhebung der Daten konnten nicht allen teilnehmenden Klassen externe Erheber zur Verfügung gestellt werden, wodurch die Objektivität der Durchführung der Verfahren nicht in allen Klassen gewährleistet werden konnte. Da aber jeder Klasse ein Lehrerheft mit standardisierten Durchführungshinweisen zur Verfügung gestellt wurde, ist davon auszugehen, dass die Verfahren weitgehend objektiv und standardisiert durchgeführt wurden.

Die *dritte Einschränkung* betrifft die Messung von Veränderungen mit dem einfachen Differenzmaß wie es in der vorliegenden Arbeit durchgeführt wurde. Wie bereits erwähnt, betrifft die Hauptkritik der Differenzmaße deren mangelnde Reliabilität. Die mangelnde Reliabilität wird zum einen durch einen doppelten Messfehler begründet, da beim Differenzmaß die Messfehler zweier Variablen von zwei Messzeitpunkten eingehen und zum anderen durch das Reliabilitäts-Validitäts-Dilemma. Das Differenzmaß scheint daher zur reliablen Abbildung von Veränderungen nicht geeignet (Rentzsch & Schütz, 2009; Rost, 1996). Bortz und Döring (2006) sowie Gollwitzer und Jäger (2009) betonen jedoch, dass Reliabilität von Differenzmaßen nicht nur von der Reliabilität der beiden Variablen zu den zwei Messzeitpunkten abhängt, sondern auch von der Autokorrelation zwischen und der Varianz zu beiden Messzeitpunkten. Ist die Autokorrelation nämlich hoch und sind die Varianzen gleich, ist die Reliabilität des Differenzmaßes niedrig, denn hohe Autokorrelation und gleiche Varianzen bedeuten, dass es keine Veränderung gab. Wenn es keine Veränderung gab, kann sie auch nicht im Differenzwert auftreten. Demzufolge sind Differenzwerte reliabel, „wenn das Messinstrument reliabel ist, die Autokorrelation kleiner als Eins ist und die Varianzen zwischen den beiden Messzeitpunkten unterschiedlich sind“ (Gollwitzer & Jäger, 2009, S. 73). Die FE-RS 2 ist durchaus reliabel und für den orthographischen Gesamtwert liegen keine Autokorrelationen vor. Da die Varianzen jedoch gleich sind, eignet sich der Differenzwert in der vorliegenden Arbeit nur eingeschränkt zur Beschreibung von Veränderungen. Eine Alternative zur Messung von Veränderungen bietet die PTT (Bortz & Döring, 2006; Klauer, 2009; Rost, 1996). Die PTT ist aufgrund der Annahmen der *Itemhomogenität* und *spezifischen Objektivität* besonders gut für Veränderungsmessungen geeignet. Die Itemhomogenität besagt, dass alle Items das gleiche Merkmal messen. Die spezifische Objektivität besagt, dass unabhängig von der Itemauswahl die Item- und Personenparameter für einen Probanden innerhalb einer modellkonformen Stichprobe immer gleich ausfallen (Fisseni, 2004; Moosbrugger, 2007). Würde ein Verfahren somit nach der PTT konstruiert werden, könnten Daten unabhängig von der Itemauswahl und der Stichprobe miteinander verglichen werden. Die Konzeption der FE-RS 2 nach der PTT sollte daher in Erwägung gezogen werden.

Trotz der methodischen Grenzen der vorliegenden Arbeit ist die FE-RS 2 ein zufriedenstellendes Verfahren, das die Rechtschreibleistungen von Schülern in der 2. Klassenstufe valide und reliabel erfasst. Um die diagnostische Gleichwertigkeit oder Überlegenheit der FE-RS 2 gegenüber gängigen Rechtschreibtests deutlich zu machen,

erfolgt in Tabelle 70 ein überblicksartiger Vergleich zwischen der FE-RS 2 und von in der deutschen Testzentrale aufgeführten Rechtschreibtests und Mehrfächertests mit einem Rechtschreibteil für die 2. Klassenstufe.

Tabelle 70: Vergleich der FE-RS 2 mit in der deutschen Testzentrale aufgeführten Rechtschreibtests und Mehrfächertests mit einem Rechtschreibteil für die 2. Klassenstufe

Verfahren	Diagnostische Zielsetzung und theoretische Grundlagen	Dauer, Anwendungshäufigkeit und Erfassung des Lernverlaufs	Auswertung	Reliabilität und Validität
FE-RS 2	<ul style="list-style-type: none"> - formative Erfassung der Rechtschreibleistung und Lernentwicklung im Rechtschreiberwerb - Kompetenzentwicklungsmodell des Rechtschreibens (Löffler & Meyer-Schepers, 2008), Entwicklungsmodell des Rechtschreibkönnens (May, 2002), Stufenmodell der Rechtschreibentwicklung (Scheerer-Neumann, 2008) 	<ul style="list-style-type: none"> - 15-45 Minuten - viermal in der 2. Klasse - <i>vorsichtige Schätzung des Lernverlaufs</i> 	quantitativ, qualitativ	<ul style="list-style-type: none"> - mittlere bis hohe Reliabilität - mittlere bis hohe Validität - <i>sehr gute Vorhersage der Rechtschreibleistung ab dem 2. Messzeitpunkt für die Form A</i>
DERET 1-2+	<ul style="list-style-type: none"> - Erfassung und Überprüfung der Rechtschreibleistungen - Stufenmodell der Schriftsprachentwicklung (Frith, 1985), Modell der zweifachen Zugangswege beim Rechtschreiben nach Diktat (Treiman, 1991), Berücksichtigung linguistischer Aspekte der Schriftsprache und orthographischer Prinzipien 	<ul style="list-style-type: none"> - 30-45 Minuten - zweimal in der 2. Klasse - keine Dokumentation des Lernverlaufs 	quantitativ, qualitativ	<ul style="list-style-type: none"> - hohe bis sehr hohe Reliabilität - mittlere bis hohe Validität - k. A. zur prognostischen Validität
DRT 2	<ul style="list-style-type: none"> - Erfassung der Rechtschreibleistung und Fehlerschwerpunkte - basiert nicht explizit auf theoretischen Annahmen des Rechtschreiberwerbs 	<ul style="list-style-type: none"> - 25-35 Minuten - einmal in der 2. Klasse - keine Dokumentation des Lernverlaufs 	quantitativ, qualitativ	<ul style="list-style-type: none"> - geringe bis sehr hohe Reliabilität und Validität - k. A. zur prognostischen Validität
HSP 1+/ HSP 2	<ul style="list-style-type: none"> - Erfassung des Rechtschreibkönnens - Fertigkeitserwerbsmodell vom Neuling zum Experten (May, 2002), Stufenmodell der Schriftsprachentwicklung (Frith, 1985), das May (2002) um die morphematische und 	<ul style="list-style-type: none"> - < 45 Minuten - zweimal in der 2. Klasse - keine Dokumentation des Lernverlaufs 	quantitativ, qualitativ	<ul style="list-style-type: none"> - hohe bis sehr hohe Reliabilität - mittlere bis hohe Validität - k. A. zur prognostischen Validität

	wortübergreifende Strategie erweiterte			
MRA	- Analyse von Rechtschreibfehlern - Modelle der Rechtschreibentwicklung, Berücksichtigung linguistischer Aspekte der Schriftsprache	- k. A. zur Dauer - zweimal in der 2. Klasse - keine Dokumentation des Lernverlaufs	quantitativ, qualitativ (onlinegestützt)	- k. A.
WRT 1+/ WRT 2+	- Erfassung der Rechtschreibfähigkeit - basiert nicht explizit auf theoretischen Annahmen des Rechtschreiberwerbs, Sichtung von Grundwortschätzen	- 45 Minuten - dreimal in der 2. Klasse - <i>vorsichtige Schätzung des Lernverlaufs</i>	quantitativ, qualitativ	- hohe bis sehr hohe Reliabilität - mittlere bis hohe Validität - k. A. zur prognostischen Validität
AST 2	- Überprüfung der spezifischen Leistungen im Rechtschreiben - basiert nicht explizit auf theoretischen Annahmen des Rechtschreiberwerbs	- 15 Minuten - einmal in der 2. Klasse - keine Dokumentation des Lernverlaufs	quantitativ	- hohe Reliabilität - mittlere Validität - k. A. zur prognostischen Validität
SLRT	- dient der „differenzierten Diagnose von Schwächen beim Erlernen des Lesens und Schreibens“ (Landerl et al., 2006, S. 7) - Zwei-Wege-Modell des Lesens und Schreibens, Annahmen über Teilkomponenten des Lesens und Rechtschreibens, Berücksichtigung linguistischer Aspekte der Schriftsprache	- 20-30 Minuten - zweimal in der 2. Klasse - keine Dokumentation des Lernverlaufs	quantitativ	- sehr geringe bis hohe Reliabilität - inhaltlich valide - k. A. zur prognostischen Validität
SBL II	- Erfassung der Rechtschreibstrategien - Stufenmodelle der Schriftsprachentwicklung, Berücksichtigung linguistischer Aspekte der Schriftsprache	- 45 Minuten - einmal in der 2. Klasse - keine Dokumentation des Lernverlaufs	quantitativ, qualitativ	- geringe bis sehr hohe Reliabilität und Validität - k. A. zur prognostischen Validität

Anmerkungen: FE-RS 2 = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr, DERET 1-2+ = Deutscher Rechtschreibtest für das erste und zweite Schuljahr (Stock & Schneider, 2008a), DRT 2 = Diagnostischer Rechtschreibtest für 2. Klassen (R. Müller, 2004), HSP 1+ = Hamburger Schreib-Probe für die Klassenstufe 1/2 (May, 2005a), HSP 2 = Hamburger Schreib-Probe für die Klasse 2 (May, 2005b), MRA = Münsteraner Rechtschreibanalyse (Schönweiss, 2008), WRT 1+ = Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen (Birkel, 2007a), WRT 2+ = Weingartener Grundwortschatz Rechtschreib-Test für zweite und dritte Klassen (Birkel, 2007b), AST 2 = Allgemeiner Schulleistungstest für zweite Klassen (Rieder & Ingenkamp, 1991), SLRT = Salzburger Lese- und Rechtschreibtest (Landerl, Wimmer & Moser, 2006), SBL II = Schultestbatterie zur Erfassung des Lernstandes in Mathematik, Lesen und Schreiben II (Kautter, Storz & Munz, 2002), k. A. = keine Angaben

Der Vergleich der FE-RS 2 mit in der Testzentrale aufgeführten Rechtschreibtests und Mehrfächertests mit einem Rechtschreibteilstest für die 2. Klassenstufe zeigt, dass die FE-RS 2 das einzige diagnostische Verfahren ist, welches die Rechtschreibleistungen zugleich frühzeitig, zeitlich engmaschig, veränderungssensitiv und theoretisch fundiert erfassen kann. Zwar erheben auch der WRT 1+ (Birkel, 2007a) und WRT 2+ (Birkel, 2007b) die Rechtschreibleistungen mehrmals innerhalb der 2. Klassenstufe, basieren aber nicht auf Theorien des Rechtschreiberwerbs. Zudem ist die FE-RS 2 das einzige diagnostische Verfahren, das als Screening konzipiert ist und Aussagen zur prognostischen Validität trifft. Dementsprechend ist es als einziges Verfahren dazu geeignet, bereits zum Ende des ersten Schulhalbjahres der 2. Klassenstufe die Schüler valide zu identifizieren, die ohne zusätzliche Förderung und differenzierende Maßnahmen im Unterricht Rechtschreibschwierigkeiten entwickeln würden. Im Vergleich zu den anderen Verfahren fallen jedoch die Werte der FE-RS 2 für die Reliabilität und Validität etwas niedriger aus, sind aber mit Werten im mittleren bis hohen Bereich für den orthographischen Gesamtwert zufriedenstellend.

Zusammenfassend betrachtet, bietet die FE-RS 2 „eine gute Möglichkeit die Rechtschreibleistungen der Schüler über das gesamte Schuljahr hinweg zu analysieren, diff. [differenzierte] Aufgaben und Förderschwerpunkte können abgeleitet werden“ (Aussage einer Lehrkraft im Lehrerfragebogen zur Einschätzung der diagnostischen Verfahren zur Lernfortschrittsmessung). Mit der FE-RS 2 liegt ein valides, reliables und sensibles Screeningverfahren vor, das

1. auf aktuellen Theorien des Rechtschreiberwerbs beruht,
2. fast ausschließlich Wörter aus den Sprachbüchern des Schulbuchkatalogs 2009/2010 für das Unterrichtsfach Deutsch der Jahrgangsstufe 2 in Grundschulen in Mecklenburg-Vorpommern (MBWK, 2009a) enthält,
3. Risikoschüler valide bestimmen kann,
4. erste vorsichtige Hinweise auf den Lernverlauf im Rechtschreiben gibt und
5. ökonomisch in der Durchführung und Auswertung ist.

Damit leistet die vorliegende Arbeit einen Beitrag zur frühen, lernprozessbegleitenden und formativen Rechtschreibdiagnostik.

Abstract

Rechtschreibschwierigkeiten können sich, wenn ihnen nicht entgegengewirkt wird, negativ auf die schulische und berufliche Zukunft auswirken. Es ist deshalb eine dringende Aufgabe, Schüler mit Schwierigkeiten im Rechtschreiberwerb so früh wie möglich zu identifizieren und zu fördern, um Rechtschreibschwierigkeiten und Folgesymptome zu verhindern (Barth & Gomm, 2008).

Für die Erfassung von Rechtschreibleistungen stehen viele normbasierter Rechtschreibtests zur Verfügung. Vor dem Hintergrund lernprozessbegleitender Diagnostik sind mit diesen diagnostischen Verfahren jedoch Probleme verbunden, da sie die Rechtschreibleistungen nicht frühzeitig, zeitlich engmaschig, veränderungssensitiv und theoretisch fundiert erfassen.

Ziel dieser Arbeit war deshalb die Entwicklung und Evaluation eines formativen Verfahrens zur Lernverlaufsdokumentation im Lernbereich Rechtschreiben der 2. Klassenstufe der Grundschule. Das Verfahren, die „Formative Erfassung der Rechtschreibleistung im 2. Schuljahr (FE-RS 2)“, sollte durch wiederholte Messungen die Lernentwicklung und den Lernverlauf im Rechtschreiberwerb dokumentieren und beurteilen können, welche Lernschritte bereits bewältigt wurden und welche nicht. Zugleich sollte die FE-RS 2 Hinweise für Planungen von Interventionen bieten und deren Evaluation ermöglichen.

Die führende Fragestellung der vorliegenden Arbeit war, ob die FE-RS 2 ein valides, reliables und veränderungssensitives Verfahren ist, das die Lernentwicklung und den Lernverlauf der Schüler im Lernbereich Rechtschreiben der 2. Klassenstufe erfassen kann.

Die FE-RS 2 wurde vor dem Hintergrund des Response to Intervention-Ansatzes, der Rechtschreibentwicklungsmodelle von Löffler und Meyer-Schepers (2008), May (2002) und Scheerer-Neumann (2008) sowie der curricularen Vorgaben für den Lernbereich Rechtschreiben in der 2. Klassenstufe (KMK, 2005; MBS et al., 2004; MBWK, 2009a) als Screeningverfahren mit zwei Parallelformen entwickelt. Das Verfahren sollte nicht alle Facetten der Rechtschreibleistung wiedergeben, sondern vorwiegend elementar graphematisches und orthographisches Können erfassen, da dieses die 2. Klassenstufe

dominiert. Darüber hinaus sollte es erste erweiterte graphematische Kenntnisse erfassen, da diese in der 2. Klassenstufe zunehmend mit dem elementaren Rechtschreibkönnen verbunden werden. Die FE-RS 2 unterscheidet demnach fünf Skalen voneinander, eine Skala für den orthographischen Gesamtwert und vier Subskalen. Die Wortwahl und das Auswertungssystem der FE-RS 2 erfolgten entsprechend entwicklungstheoretisch, linguistisch fundiert und curriculumorientiert.

Die Untersuchung im Cross-Over-Design umfasste, im Abstand von jeweils 10 Wochen, 4 Messzeitpunkte. Die Untersuchung wurde in 29 2. Klassen (N = 644) von Grundschulen und Regionalen Schulen mit Grundschulteil im Schulamtsbezirk Neubrandenburg in Mecklenburg-Vorpommern von November 2009 bis Juni 2010 durchgeführt.

Die in der Untersuchung ermittelten Ergebnisse zeigen, dass die FE-RS 2, vorwiegend der orthographische Gesamtwert der Form A, die Rechtschreibleistungen von Zweitklässlern objektiv, reliabel und valide erfassen kann. Die Ergebnisse zeigen zudem, dass der orthographische Gesamtwert der Form A prognostische Validität besitzt und zum Ende des ersten Schulhalbjahres der 2. Klassenstufe Schüler mit zukünftigen Rechtschreibschwierigkeiten richtig voraussagen kann. Die Annahme der Äquivalenz der Formen A und B der FE-RS 2 kann aufgrund der ermittelten Ergebnisse nicht bestätigt werden. Somit liegen keine Parallelförmigkeiten vor. Auch die Annahme einer validen und reliablen Lernverlaufs- und Veränderungsmessungen mit der FE-RS 2 kann nicht bestätigt werden, da sich Differenzmaße nicht zur Beschreibung von Veränderungen innerhalb der klassischen Testtheorie eignen.

Zusammenfassend betrachtet, ist die FE-RS 2 ein zufriedenstellendes Screeningverfahren, das die Rechtschreibleistungen theoriegeleitet und curriculumbasiert erfasst. Die FE-RS 2 ermöglicht, Unterricht zu evaluieren und pädagogische Maßnahmen zu planen.

Literaturverzeichnis

- Ahlgrimm, H., Bartkowski, K., Karasz, C. & Pfeiffer, S. (2004). *Bücherwurm Sprachbuch* 2. Leipzig: Klett.
- AIMSweb. (2003). *AIMSweb® Standard Spelling Benchmark Assessment List #1 (3rd Grade)*. Zugriff am 22.09.2010. Verfügbar unter http://www.aimsweb.com/uploads/pdfs/sample_scbm.pdf
- Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung* (6., vollständig überarbeitete Aufl.). Stuttgart: Kohlhammer.
- Arnold, K.-H. & Hartig, J. (2008). Testtheoretische Grundlagen der Förderdiagnostik. In K.-H. Arnold, O. Graumann & A. Rakhokhina (Hrsg.), *Handbuch Förderung* (S. 116-125). Weinheim: Beltz.
- Augst, G. & Dehn, M. (2007). *Rechtschreibung und Rechtschreibunterricht. Können – Lehren – Lernen. Eine Einführung für Studierende und Lehrende aller Schulformen* (3., überarbeitete und aktualisierte Aufl.). Stuttgart: Klett.
- Barry, C. (1994). Spelling Routes (or Roots or Rutes). In G. D. A. Brown & N. C. Ellis (Hrsg.), *Handbook of Spelling. Theory, Process and Intervention* (S. 27-49). Chichester: John Wiley & Sons.
- Barth, K. & Gomm, B. (2008). Gruppentest zur Früherkennung von Lese- und Rechtschreibschwierigkeiten. Phonologische Bewusstheit bei Kindergartenkindern und Schulanfängern (PB-LRS). In W. Schneider, H. Marx & M. Hasselhorn (Hrsg.), *Diagnostik von Rechtschreibleistungen und -kompetenz* (S. 7-43).
- Bauch, B., Dirzus, U., Greven, M., Hinze, G., Isack, A. & Schmidt, H.-P. (2008). *Bausteine Sprachbuch 2*. Braunschweig: Diesterweg.
- Baumann, D., Pristl, T. & Schmidt, J. (2004). *Das Sprachbuch 2* (Ausgabe D). München: bsv.
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., Müller, P. et al. (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.
- Birkel, P. (unter Mitarbeit von Stammet, C.) (2007a). *WRT 1+. Weingartener Grundwortschatz. Rechtschreib-Test für erste und zweite Klassen* (2., neu normierte und vollständig überarbeitete Aufl.). Göttingen: Hogrefe.
- Birkel, P. (unter Mitarbeit von Stammet, C.) (2007b). *WRT 2+. Weingartener Grundwortschatz. Rechtschreib-Test für zweite und dritte Klassen* (2., neu normierte und vollständig überarbeitete Aufl.). Göttingen: Hogrefe.
- Blankenship, C. S. (1985). Using Curriculum-Based Assessment Data to Make Instructional Decisions. *Exceptional Children*, 52, 233-238.

- Böhme, K. & Bremerich-Vos, A. (2009). Diagnostik der Rechtschreibkompetenz in der Grundschule – Konstruktprüfung mittels Fehler- und Dimensionsanalysen. In D. Granzer, O. Köller, A. Bremerich-Vos, A. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 330-356). Weinheim: Beltz.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6., vollständig überarbeitete und erweiterte Aufl.). Berlin: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4., überarbeitete Aufl.). Heidelberg: Springer.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., vollständig überarbeitete und erweiterte Aufl.). Berlin: Springer.
- Bradley, R., Danielson, L. & Doolittle, J. (2007). Responsiveness to Intervention: 1997 to 2007. *Teaching exceptional children*, 39 (5), 8-12.
- Braun, O. (2006). *Sprachstörungen bei Kindern und Jugendlichen. Diagnostik – Therapie – Förderung* (3., aktualisierte und erweiterte Aufl.). Stuttgart: Kohlhammer.
- Brosius, F. (2011). *SPSS 19*. Heidelberg: mitp.
- Brown, G. D. A. & Loosemore, R. P. W. (1994). Computational Approaches to Normal and Impaired Spelling. In G. D. A. Brown & N. C. Ellis (Hrsg.), *Handbook of Spelling. Theory, Process and Intervention* (S. 319-335). Chichester: John Wiley & Sons.
- Brügelmann, H. (1986). Fehler: »Defekte« im Leistungssystem oder individuelle Annäherungsversuche an einen schwierigen Gegenstand? In H. Brügelmann (Hrsg.), *ABC und Schriftsprache: Rätsel für Kinder, Lehrer und Forscher* (S. 22-31). Konstanz: Faude.
- Bühl, A. (2008). *SPSS 16. Einführung in die moderne Datenanalyse* (11., aktualisierte Aufl.). München: Pearson Studium.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson Studium.
- Bundschuh, K. (2005). *Einführung in die sonderpädagogische Diagnostik* (6., aktualisierte Aufl.). München: Reinhardt.
- Burgel, A., Kelch, S., Meeh, S., Naumann-Harms, H., Rahm, K., Stanzel, R. et al. (2004). *Jo-Jo Sprachbuch 2. Ausgabe C*. Berlin: Cornelsen.
- Bußmann, H. (2008). *Lexikon der Sprachwissenschaft* (4., durchgesehene und bibliographisch ergänzte Aufl.). Stuttgart: Kröner.
- Capizzi, A. M. & L. S. Fuchs (2005). Effects of Curriculum-Based Measurement With and

- Without Diagnostic Feedback on Teacher Planning. *Remedial and Special Education*, 26, 159-174.
- Caplan, G. (1964). *Principles of preventive psychiatry*. New York, NY: Basic Books.
- Caspi, O. & Bell, I. R. (2004). One Size Does Not Fit All: Aptitude x Treatment Interaction (ATI) as a Conceptual Framework for Complementary and Alternative Medicine Outcome Research. Part 1-What Is ATI Research? *The Journal of Alternative and Complementary Medicine*, 10, 580-586.
- Cattel, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1 (CFT 1)* (5. Aufl.). Braunschweig: Westermann.
- Cholewa, J. (2004, März). Analyse von Schreibfehlern auf psycholinguistischer Grundlage. *Osnabrücker Beiträge zur Sprachtheorie*, (67), 119-141.
- Christ, T. J., Burns, M. K. & Ysseldyke, J. E. (2005). Conceptual Confusion Within Response-to-Intervention Vernacular: Clarifying Meaningful Differences. *Communiqué*, 34 (3), o. S.. Zugriff am 03.03.2009. Verfügbar unter <http://www.nasponline.org/publications/cq/cq343rti.aspx>
- Clasen, H. (2010). *Die Messung von Lernerfolg: Eine grundsätzliche Aufgabe der Evaluation von Lehr- bzw. Trainingsinterventionen*. Dissertation, Technische Universität Dresden. Zugriff am 16.02.2011. Verfügbar unter http://www.qucosa.de/fileadmin/data/qucosa/documents/2700/Dissertation_HeidiClasen.pdf
- Corno, L., Cronbach, L. J., Kupermintz, H., Lohman, D. F., Mandinach, E. B., Porteus, A. W. et al. (2002). *Remaking the Concept of Aptitude: Extending the Legacy of Richard E. Snow*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Costard, S. (2007). *Störungen der Schriftsprache. Modellgeleitete Diagnostik und Therapie*. Stuttgart: Thieme.
- Cronbach, L. J. (1957). The Two Disciplines of Scientific Psychology. *The American Psychologist*, 12, 671-684.
- Cronbach, L. J. (1967). How Can Instruction be Adapted to Individual Differences? In R. M. Gagné (Hrsg.), *Learning and Individual Differences* (S. 23-39). Columbus, OH: Merrill.
- Cronbach, L. J. (1975a). Beyond the Two Disciplines of Scientific Psychology. *The American Psychologist*, 30, 116-127.
- Cronbach, L. J. (1975b). Wie kann Unterricht an individuelle Unterschiede angepaßt werden? In R. Schwarzer & K. Steinhagen (Hrsg.), *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (S. 42-58). München: Kösel.
- Cronbach, L. J. & Snow, R. E. (1969). *Final Report. Individual Differences in Learning Ability as a Function of Instructional Variables*. Stanford, CA: Universität.

- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods. A handbook for research on interactions*. New York, NY: Irvington Publishers.
- Daly III, E. J., Martens, B. K., Barnett, D., Witt, J. C. & Olson, S. C. (2007). Varying Intervention Delivery in Response to Intervention: Confronting and Resolving Challenges With Measurement, Instruction, and Intensity. *School Psychology Review*, 36, 562-581.
- Daneš, F. (1982). Zur Theorie des sprachlichen Zeichensystems. In J. Scharnhorst & E. Ising (Hrsg.), *Grundlagen der Sprachkultur. Beiträge der Prager Linguistik zur Sprachtheorie und Sprachpflege* (Teil 2). Berlin (DDR): Akademie-Verlag.
- Deno, S. L. (2003). Developments in Curriculum-Based Measurement. *The Journal of Special Education*, 37, 184-192.
- Deno, S. L., Mirkin, P. K., Lowry, L. & Kuehnle, K. (1980). *Relationships among simple measures of spelling and performance on standardized achievement tests* (Research Report No. 21). Minneapolis, MN: University, Institute for Research on Learning Disabilities.
- Diehl, K. (2008). *Inventar zur Erfassung der Lesekompetenzen von Erstklässlern (IEL-1) - theoretische Grundlagen, Konzeption, Güte des Verfahrens*. Dissertation, Universität Rostock.
- Diehl, K. & Hartke, B. (2007). Curriculumnahe Lernfortschrittsmessungen. *Sonderpädagogik*, 37, 195-211.
- Diehl, K. & Hartke, B. (2012). *Inventar zur Erfassung der Lesekompetenzen im 1. Schuljahr (IEL-1). Ein curriculumbasiertes Verfahren zur Abbildung des Lernfortschritts*. Göttingen: Hogrefe.
- Dinges, E. (2005). *Systematische Beurteilung und Förderung schulischer Leistungen* (2. Aufl.). Horneburg: Persen.
- Dolenc, R., Fiskus, C., Kraft, G., Röbe, E. & Röbe, H. (2007). *Das Auer Sprachbuch. Schulbuch für die 1./2. Klasse* (Ausgabe N). Donauwörth: Auer.
- Dreyfus, H. L. & Dreyfus, S. E. (1991). *Künstliche Intelligenz. Von den Grenzen der Denkmachine und dem Wert der Intuition*. Reinbek bei Hamburg: Rowohlt.
- Dürscheid, C. (2006). *Einführung in die Schriftlinguistik* (3., überarbeitete und ergänzte Aufl.). Göttingen: Vandenhoeck & Ruprecht.
- Eichler, W. (1986). Zu Uta Frith' Dreiphasenmodell des Lesen (und Schreiben) Lernens. Oder: Lassen sich verschiedene Modelle des Schrifterwerbs aufeinander beziehen und weiterentwickeln? On Uta Frith's three phase model of learning to read and to write. Or: Can different models of the acquisition of reading and writing be interrelated and further developed? In G. Augst (Hrsg.), *New Trends in Graphemics and Orthography* (S. 234-247). Berlin: de Gruyter.
- Eisenberg, P. (1996). Sprachsystem und Schriftsystem. In H. Günther & O. Ludwig

- (Hrsg.), *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 10, Halbbd. 2, S. 1368-1380). Berlin: de Gruyter.
- Eisenberg, P. (2006). *Grundriss der deutschen Grammatik. Das Wort* (Bd. 1, 3., durchgesehene Aufl.). Stuttgart: Metzler.
- Enderle, S. (2005). *Autonomie der geschriebenen Sprache? Zur Theorie phonographischer Beschreibungskategorien am Beispiel des Deutschen*. Berlin: Erich Schmidt Verlag.
- Farrington-Flint, L., Stash, A. & Stiller, J. (2008). Monitoring variability and change in children's spelling strategies. *Educational Psychology*, 28, 133-149.
- Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik. Mit Hinweisen zur Intervention* (3., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Fletcher, J. M. (2008). *Identifying Learning Disabilities in the Context of Response to Intervention: A Hybrid Model*. Zugriff am 02.03.2009. Verfügbar unter: <http://www.rtinetwork.org/Learn/LD/ar/HybridModel>
- Folk, J. R., Rapp, B. & Goldrick, M. (2002). The interaction of lexical and sublexical information in spelling: What's the point? *Cognitive Neuropsychology*, 19, 653-671.
- Friedrich, B. (1996). Aspekte und Probleme des Schreibunterrichts: Rechtschreiben. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 10, Halbbd. 2, S. 1249-1260). Berlin: de Gruyter.
- Frith, U. (1985). Beneath the Surface of Developmental Dyslexia. In K. E. Patterson, J. C. Marshall & M. Coltheart (Hrsg.), *Surface Dyslexia. Neuropsychological and Cognitive Studies of Phonological Reading* (S. 301-330). London: Erlbaum.
- Frith, U. (1986). Psychologische Aspekte des orthographischen Wissens. Entwicklung und Entwicklungsstörung. Psychological aspects of orthographic skills: development and disorder. In G. Augst (Hrsg.), *New Trends in Graphemics and Orthography* (S. 218-233). Berlin: de Gruyter.
- Fuchs, D. & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41, 93-99.
- Fuchs, D., Mock, D., Morgan, P. L. & Young, C. L. (2003). Responsiveness-to-Intervention: Definitions, Evidence, and Implications for the Learning Disabilities Construct. *Learning Disabilities Research & Practice*, 18, 157-171.
- Fuch, L. S. & Deno, S. L. (1991). Paradigmatic Distinctions between Instructionally Relevant Measurement Models. *Exceptional Children*, 57, 488-499.
- Fuchs, L. S. & Fuchs, D. (2007). A Model for Implementing Responsiveness to Intervention. *Teaching exceptional children*, 39 (5), 14-20.
- Fuchs, L. S. & Fuchs, D. (n. d.). *Using curriculum-based measurement for progress*

monitoring in math. Zugriff am 10.09.2010. Verfügbar unter http://www.k12.wa.us/SpecialEd/pubdocs/RTI/CBM_Manual_Math_9_29_05.pdf

Gage, N. L. & Berliner, D. C. (1996). *Pädagogische Psychologie* (5., vollständig überarbeitete Aufl.). Weinheim: Beltz.

Gallmann, P. (1985). *Graphische Elemente der geschriebenen Sprache. Grundlagen für eine Reform der Orthographie*. Tübingen: Niemeyer.

Gallmann, P. & Sitta, H. (1998). *Handbuch Rechtschreiben* (4., unveränderte Aufl.). Zürich: Lehrmittelverlag des Kantons Zürich.

Garten, H.-K. (1980). Zur inhaltlichen Problematik von Aptitude-Treatment-Interaktion. *Zeitschrift für erziehungswissenschaftliche Forschung*, 14 (1), 13-35.

Gasteiger-Klicpera, B. & Klicpera, C. (2004). Lese-Rechtschreib-Schwäche. In G. W. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen. Förderung, Training und Therapie in der Praxis* (S. 46-54). Göttingen: Hogrefe.

Gasteiger-Klicpera, B., Klicpera, C. & Schabmann, A. (2006). Der Zusammenhang zwischen Lese-, Rechtschreib- und Verhaltensschwierigkeiten. Entwicklung vom Kindergarten bis zur vierten Grundschulklasse. *Kindheit und Entwicklung*, 15, 55-67.

Gickling, E. E. & Havertape, J. (1981). Curriculum Based Assessment (CBA). In J. A. Tucker (Hrsg.), *Non Test-Based Assessment. Trainer Manual* (S. 212-497). Minneapolis, MN: The National School Psychology Inservice Training Network. Zugriff am 15.09.2010. Verfügbar unter <http://www.eric.ed.gov/PDFS/ED236864.pdf>

Glaser, R. (1972). Individuals and Learning: The New Aptitudes. *Educational Researcher*, 1 (6), 5-13.

Glaser, R. (1975). Individuen und Lernen: Die neuen Schülermerkmale. In R. Schwarzer & K. Steinhagen (Hrsg.), *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (S. 109-126). München: Kösel.

Glück, H. (unter Mitarbeit von Schmöe, F.) (2005). *Metzler Lexikon Sprache* (3., neubearbeitete Aufl.). Stuttgart: Metzler.

Goethe-Institut Inter Nationes. (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Zugriff am 21.09.2010. Verfügbar unter <http://www.goethe.de/z/50/commeuro/i0.htm>

Goldhammer, F. & Hartig, J. (2007). Interpretation von Testresultaten und Testeichung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 165-192). Heidelberg: Springer.

Gollwitzer, M. & Jäger, R. S. (2009). *Evaluation kompakt*. Weinheim: Beltz.

Grassegger, H. (2010). *Phonetik/ Phonologie* (4. Aufl.). Idstein: Schulz-Kirchner.

Gresham, F. M. (2007). Evolution of the Response-to-Intervention Concept: Empirical

- Foundations and Recent Developments. In S. R. Jimerson, M. K. Burns & A. M. VanDerHeyden (Hrsg.), *Handbook of Response to Intervention. The Science and Practice of Assessment and Intervention* (S. 10-24). New York, NY: Springer.
- Gresham, F. M., VanDerHeyden, A. & Witt, J. C. (2005). *Response to Intervention in the Identification of Learning Disabilities: Empirical Support and Future Challenges*. Zugriff am 03.03.2009. Verfügbar unter <http://www.joewitt.org/Downloads/Response%20to%20Intervention%20MS%20Gresham%20%20Vanderheyden%20Witt.pdf>
- Grube, D. & Hasselhorn, M. (2006). Längsschnittliche Analysen zur Lese-, Rechtschreib- und Mathematikleistung im Grundschulalter: zur Rolle von Vorwissen, Intelligenz, phonologischem Arbeitsgedächtnis und phonologischer Bewusstheit. In I. Hosenfeld & F.-W. Schrader (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven* (S. 87-105). Münster: Waxmann.
- Grund, M., Haug, G. & Naumann, C. L. (2004a). *DRT 4. Diagnostischer Rechtschreibtest für 4. Klassen* (2., aktualisierte Aufl.). Göttingen: Beltz Test.
- Grund, M., Haug, G. & Naumann, C. L. (2004b). *DRT 5. Diagnostischer Rechtschreibtest für 5. Klassen* (2., aktualisierte Aufl.). Göttingen: Beltz Test.
- Günther, K. B. (1986). Ein Stufenmodell der Entwicklung kindlicher Lese- und Schreibstrategien. In H. Brügelmann (Hrsg.), *ABC und Schriftsprache: Rätsel für Kinder, Lehrer und Forscher* (S. 32-54). Konstanz: Faude.
- Hale, J. B., Kaufman, A., Naglieri, J. A. & Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools*, 43, 753-770.
- Hamisch, S., Hofmann, G., Hübner, A., Mühlbauer, E., Rehak, B., Scheida, D. et al. (2005). *Duden Sprachbuch 2*. Berlin: Duden Paetec.
- Hartmann, E. (2008). Konzeption und Diagnostik von schriftsprachlichen Lernstörungen im Responsiveness-to-Intervention-Modell: eine kritische Würdigung. *Vierteljahreszeitschrift für Heilpädagogik und ihre Nachbargebiete*, 77, 123-137.
- Hartmann, E. & Müller, C. (2009). Schulweite Prävention von Lernproblemen im RTI-Modell. *Schweizerische Zeitschrift für Heilpädagogik*, 15 (9), 25-33.
- Hasebrook, J. (2006). Aptitude-Treatment-Interaktion. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (3., überarbeitete und erweiterte Aufl., S. 20-26). Weinheim: Beltz.
- Heller, K. A., Holtzman, W. H. & Messick, S. (1982). *Placing children in special education: A strategy for equity*. Washington, DC: National Academy Press.
- Helmke, A. (2007). Guter Unterricht – nur ein Angebot? Interview mit dem Unterrichtsforscher Andreas Helmke. *Friedrich Jahresheft*, 25, 62-63.
- Hemminger, U., Roth, E., Schneck, S., Jans, T. & Warnke, A. (2000). Testdiagnostische

Verfahren zur Überprüfung der Fertigkeiten im Lesen, Rechtschreiben und Rechnen. Eine kritische Übersicht. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 28, 188-201.

Herné, K.-L. (2004, März). Diagnose der Rechtschreibfähigkeiten. *Osnabrücker Beiträge zur Sprachtherapie*, (67), 89-113.

Herné, K.-L. (2006). Rechtschreibtests. In U. Bredel, H. Günther, P. Klotz, J. Ossner & G. Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache. Ein Handbuch* (Bd. 2, 2., durchgesehene Aufl., S. 883-897). Paderborn: Schöningh.

Herné, K.-L. & Naumann, C. L. (2002). *Aachener Förderdiagnostische Rechtschreibfehler-Analyse – AFRA. Systematische Einführung in die Praxis der Fehleranalyse mit Auswertungshilfen zu insgesamt 33 standardisierten Testverfahren als Kopiervorlagen mit Beiträgen von Cordula Löffler*. Aachen: Alfa Zentaurus.

Hinney, G., Huneke, H.-W., Müller, A. & Weinhold, S. (2008, Sonderheft). Definition und Messung von Rechtschreibkompetenz. *Didaktik Deutsch*, (2), 107-126.

Hoover, J. J., Baca, L., Wexler-Love, E. & Saenz, L. (2008). *National Implementation of Response to Intervention (RTI): Research Summary*. Boulder, CO: University of Colorado-Boulder. Zugriff am 08.09.2010. Verfügbar unter <http://nasdse.org/Portals/0/NationalImplementationofRTI-ResearchSummary.pdf>

Hosp, M. K., Hosp, J. L. & Howell, K. W. (2007). *The ABCs of CBM. A Practical Guide to Curriculum-Based Measurement*. New York, NY: The Guilford Press.

Howell, K. W., Kurns, S. & Antil, L. (2002). Best Practices in Curriculum-Based Evaluation. In A. Thomas & J. Grimes (Hrsg.), *Best Practices in School Psychology IV* (Bd. 1, 4. Aufl., S. 753-771). Bethesda, MD: National Association of School Psychologists.

Ingenkamp, K. & Lissmann, U. (2008). *Lehrbuch der Pädagogischen Diagnostik* (6., neu ausgestattete Aufl.). Weinheim: Beltz.

Jaeschke, R., Guyatt, G. & Sackett, D. L. (1994). Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA*, 271, 703-707. Zugriff am 04.04.2012. Verfügbar unter http://www.hopkinsmedicine.org/gim/_pdf/JAMA/3B-Diagnostic_Test.pdf

Jansen, H. (2008). Rechtschreibdiagnostik in pädagogisch-psychologischen Interventionen: Leistungen und Grenzen heutiger Rechtschreibtests in der praktischen Anwendung. In W. Schneider, H. Marx & M. Hasselhorn (Hrsg.), *Diagnostik von Rechtschreibleistungen und -kompetenz* (Jahrbuch der pädagogisch-psychologischen Diagnostik, Tests und Trends, Neue Folge Bd. 6, S. 159-189). Göttingen: Hogrefe.

Jansen, H., Mannhaupt, G. & Marx, H. (1993). Probleme bei der Übertragbarkeit anglo-amerikanischer Entwicklungsmodelle des Lesens und Rechtschreibens auf deutschsprachige Schulkinder. Empirische Untersuchung zur Existenz der "logographischen Stufe" des Lesenlernens bei Erstkläßlern. In H. Bauersfeld & R.

- Bromme (Hrsg.), *Bildung und Aufklärung. Studien zur Rationalität des Lehrens und Lernens. Festschrift für Helmut Skowronek zum 60. Geburtstag* (S. 69-88). Münster: Waxmann.
- Jogschies, P. (2008). Diskussion grundlegender Begriffe. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), *Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher* (S. 19-36). Stuttgart: Kohlhammer.
- Kallós, D. (1975). Zur Bedeutung der ATI-Forschung für die Pädagogik und für die Theorie und Praxis des Unterrichts. In R. Schwarzer & K. Steinhagen (Hrsg.), *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (S. 174-186). München: Kösel.
- Kaminski, R., Cummings, K. D., Powell-Smith, K. A. & Good III, R. H. (2008). Best practices in using Dynamic Indicators Basic Early Literacy Skills (DIBELS) for formative assessment and evaluation. In A. Thomas & J. Grimes (Hrsg.), *Best practices in school psychology V* (Bd. 4, 6. Aufl., S. 1181-1204). Bethesda, MD: National Association of School Psychologists.
- Kautter, H., Storz, L. & Munz, W. (2002). *SBL II. Schultestbatterie zur Erfassung des Lernstandes in Mathematik, Lesen und Schreiben II*. Göttingen: Beltz Test.
- Kelava, A. & Moosbrugger, H. (2007). Deskriptivstatistische Evaluation von Items (Itemanalyse) und Testwertverteilungen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 73-98). Heidelberg: Springer.
- Kemmler, L. (1976). *Schulerfolg und Schulversagen. Eine Längsschnittuntersuchung vom ersten bis zum fünfzehnten Schulbesuchsjahr*. Göttingen: Hogrefe.
- Kirschhock, E.-M. (2004). *Entwicklung schriftsprachlicher Kompetenzen im Anfangsunterricht*. Bad Heilbrunn/Obb.: Klinkhardt.
- Klauer, K. J. (2006). Erfassung des Lernfortschritts durch curriculumbasierte Messung. *Heilpädagogische Forschung*, 32, 16-26.
- Klauer, K. J. (2009). Veränderungsmessung. In H. Holling (Hrsg.), *Grundlagen und statistische Methoden der Evaluationsforschung* (S. 525-563). Göttingen: Hogrefe.
- Klauer, K. J. (2011). Lernverlaufsdiagnostik – Konzept, Schwierigkeiten und Möglichkeiten. *Empirische Sonderpädagogik*, 3, 207-224.
- Klauer, K. J. & Leutner, D. (2007). *Lehren und Lernen. Einführung in die Instruktionspsychologie*. Weinheim: Beltz PVU.
- Klicpera, C., Gasteiger-Klicpera, B. & Schabmann, A. (1993). *Lesen und Schreiben. Entwicklung und Schwierigkeiten. Die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit*. Bern: Huber.
- Klicpera, C., Schabmann, A., Gasteiger-Klicpera, B. (2006). Die mittelfristige Entwicklung von Schülern mit Teilleistungsschwierigkeiten im Bereich der Lese- und

Rechtschreibschwierigkeiten. *Kindheit und Entwicklung*, 15, 216-227.

- Klicpera, C., Schabmann, A., Gasteiger-Klicpera, B. (2007). *Legasthenie. Modelle, Diagnose, Therapie und Förderung* (2., aktualisierte Aufl.). München: Reinhardt.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2007). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise* (Bildungsforschung, Bd. 1). Bonn: Bundesministerium für Bildung und Forschung.
- Knopp, E. (2010). *Theoretische Grundlagen, Konzeption und Güte des Inventars „Rechenfische“. Ein Verfahren zur Dokumentation von Fortschritten beim Erlernen arithmetischer Kenntnisse im Anfangsunterricht Mathematik*. Dissertation, Universität Rostock.
- Knopp, E. & Hartke, B. (2010). Das Inventar Rechenfische – Anwendung, Reliabilität und Validität eines Verfahrens zur Erfassung des Leistungsstandes von Erstklässlern in Mathematik. *Empirische Sonderpädagogik*, 2, 5-25.
- Kosog, O. (1912). *Unsere Rechtschreibung und die Notwendigkeit ihrer gründlichen Reform*. Leipzig: Teubner.
- Krajewski, K. (2008). Vorschulische Förderung bei beeinträchtigter Entwicklung mathematischer Kompetenzen. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), *Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher* (S. 122-135). Stuttgart: Kohlhammer.
- Krajewski, K., Küspert, P. & Schneider, W. (2002). *Deutscher Mathematiktest für erste Klassen (DEMAT 1+)*. Göttingen: Beltz.
- Krajewski, K., Liehm, S. & Schneider, W. (2004). *Deutscher Mathematiktest für zweite Klassen (DEMAT 2+)*. Göttingen: Beltz.
- Krapp, A. (1984). Forschungsergebnisse zur Bedingungsstruktur der Schulleistung. In K. A. Heller (Hrsg.), *Leistungsdiagnostik in der Schule* (4., völlig neubearbeitete Aufl., S. 46-62). Bern: Huber.
- Kretschmann, R. (2000). Präventionsmodelle in der Schule. In J. Borchert (Hrsg.), *Handbuch der Sonderpädagogischen Psychologie* (S. 325-340). Göttingen: Hogrefe.
- Kretschmann, R. (2003). Prävention und pädagogische Interventionen bei Beeinträchtigungen des Lernens. In A. Leonhardt & F. B. Wember (Hrsg.), *Grundfragen der Sonderpädagogik. Bildung, Erziehung, Behinderung* (S. 465-503). Weinheim: Beltz.
- Kuhlmann, K. & Hartke, B. (2009a). *Formative Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr (FE-AF 2)*. Rostock: Universität, Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation.
- Kuhlmann, K. & Hartke, B. (2009b). *Formative Erfassung der Lesefertigkeit im 2. Schuljahr (FE-L 2)*. Rostock: Universität, Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation.

- Küspert, P. & Schneider, W. (1998). *Würzburger Leise Leseprobe (WLLP)*. Göttingen: Hogrefe.
- Landerl, K., Wimmer, H. & Moser, E. (2006). *SLRT. Salzburger Lese- und Rechtschreibtest. Verfahren zur Differentialdiagnose von Störungen der Teilkomponenten des Lesens und Schreibens für die 1. bis 4. Schulstufe* (2., korrigierte und aktualisierte Aufl.). Bern: Huber.
- Landesinstitut für Schule und Medien Brandenburg (2006). *Hinweise zur Umsetzung der KMK-Bildungsstandards für den Primarbereich (Jahrgangsstufe 4). Materialien zu den KMK-Bildungsstandards*. Zugriff am 06.08.2010. Verfügbar unter http://bildungsserver.berlin-brandenburg.de/fileadmin/bbs/unterricht_und_pruefungen/bildungsstandards/01.02.06_Hinweise_Primarstufe_3.0_01.pdf
- Lenhard, W. (2011). *Auswertungshilfe zur Berechnung von Testkennwerten von Screeningverfahren*. Zugriff am 24.03.2012. Verfügbar unter <http://www.psychometrica.de/testkennwerte.html>
- Leonhart, R. (2009). *Lehrbuch Statistik. Einstieg und Vertiefung* (2., überarbeitete und erweiterte Aufl.). Bern: Huber.
- Leutner, D. (1992). *Adaptive Lehrsysteme. Instruktionspsychologische Grundlagen und experimentelle Analysen*. Weinheim: Psychologie Verlags Union.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Beltz.
- Lindquist, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education*. Boston, MA: Houghton Mifflin Company.
- L.I.S.A. Mecklenburg-Vorpommern (2006). *Glossar. -Ausgewählte pädagogische Begriffe-*. Zugriff am 06.08.2010. Verfügbar unter http://www.bildung-mv.de/export/sites/lisa/de/publikationen/rahmenplaene/ergaenzende_texte/Glossar_paed_Begriffe.pdf
- Lischeid, T. (2007). *Rechtschreiben. Ringvorlesung „Deutschunterricht heute“*. Zugriff am 22.07.2010. Verfügbar unter http://www.ruhr-uni-bochum.de/lidi/RVL_Du_heute_SoSe_07_RechtschreibenLisch.pdf
- Löffler, I. & Meyer-Schepers, U. (1992). *DoRA. Dortmunder Rechtschreibfehler Analyse zur Ermittlung des Schriftsprachstatus rechtschreibschwacher Schüler*. Dortmund: ILT.
- Löffler, I. & Meyer-Schepers, U. (2005). Orthographische Kompetenzen: Ergebnisse qualitativer Fehleranalysen, insbesondere bei schwachen Rechtschreibern. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 81-108). Münster: Waxmann.
- Löffler, I. & Meyer-Schepers, U. (2006). Probleme beim Erwerb von Rechtschreibkompetenz: Ergebnisse qualitativer Fehleranalyse aus IGLU-E. In S. Weinhold (Hrsg.), *Schriftspracherwerb empirisch. Konzepte – Diagnostik –*

- Entwicklung* (S. 199-217). Baltmannsweiler: Schneider Verlag Hohengehren.
- Löffler, I. & Meyer-Schepers, U. (2007). Beschreibung von Rechtschreibschwächen mit einem theoretisch fundierten Kompetenzmodell. In B. Hofmann & R. Valtin (Hrsg.), *Förderdiagnostik beim Schriftspracherwerb* (S. 179-196). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Löffler, I. & Meyer-Schepers, U. (2008, Juli). Analyse von Rechtschreibfehlern. Diagnose mit Hilfe von Kompetenzmodellen. *Deutsch differenziert*, (3), 30-32.
- Löffler, I., Meyer-Schepers, U. & Lischeid, T. (2007). Rechtschreibschwäche im Fokus der Kompetenzdiagnostik. Ergebnisse aus der internationalen Grundschul-Lese-Untersuchung (IGLU-E). In G. Schulte-Körne (Hrsg.) (in Zusammenarbeit mit dem Bundesverband Legasthenie und Dyskalkulie e.V.), *Legasthenie und Dyskalkulie: Aktuelle Entwicklungen in Wissenschaft, Schule und Gesellschaft* (S. 267-280). Bochum: Winkler.
- Löffler, I., Meyer-Schepers, U. & Stubbe, T. C. (2008). Orthografiekompetenzen von Viertklässlerinnen und Viertklässlern in der Deutschsprachigen Gemeinschaft. In W. Bos, S. Sereni & T. C. Stubbe (Hrsg.), *Lese- und Orthografiekompetenzen von Grundschulkindern in der Deutschsprachigen Gemeinschaft* (S. 137-145). Münster: Waxmann.
- Ludwig, O. (1983). Einige Vorschläge zur Begrifflichkeit und Terminologie von Untersuchungen im Bereich der Schriftlichkeit. In K. B. Günther & H. Günther (Hrsg.), *Schrift, Schreiben, Schriftlichkeit. Arbeiten zur Struktur, Funktion und Entwicklung schriftlicher Sprache* (S. 1-15). Tübingen: Niemeyer.
- Maas, U. (1991). Die Rechtschreibung als wissensbasiertes System. *Osnabrücker Beiträge zur Sprachtheorie*, (44), 13-39.
- Mannhaupt, G. (2001). *Lernvoraussetzungen im Schriftspracherwerb. Eine Studie zur Entwicklung der Schriftsprach- und ihrer Teilfertigkeiten sowie deren Voraussetzungen im Vor- und Grundschulalter*. Köln: Kölner Studien Verlag.
- Mannhaupt, G. (2008). Prävention von Lese-Rechtschreibschwierigkeiten im Kindergarten. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), *Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher* (S. 136-148). Stuttgart: Kohlhammer.
- Mannhaupt, G. (2010). Phonologische Bewusstheit – eine notwendige Voraussetzung des Schriftspracherwerbs. In B. Hartke, K. Koch & K. Diehl (Hrsg.), *Förderung in der schulischen Eingangsstufe* (S. 186-209). Stuttgart: Kohlhammer.
- Marston, D. (1982). *The technical adequacy of direct, repeated measurement of academic skills in low-achieving elementary students*. Unpublished doctoral dissertation, University of Minneapolis, MN.
- Marston, D. B. (1989). A Curriculum-Based Measurement Approach to Assessing Academic Performance: What It Is and Why Do It. In M. R. Shinn (Hrsg.), *Curriculum-Based Measurement. Assessing Special Children* (S. 18-78). New York, NY: The Guilford Press.

- Marx, H., Jansen, H. & Skowronek, H. (2000). Prognostische, differentielle und konkurrente Validität des Bielefelder Screenings zur Früherkennung von Lese-Rechtschreibschwierigkeiten. In M. Hasselhorn, W. Schneider & H. Marx (Hrsg.), *Diagnostik von Lese-Rechtschreibschwierigkeiten* (Jahrbuch der pädagogisch-psychologischen Diagnostik, Tests und Trends, Neue Folge Bd. 1, S. 9-34). Göttingen: Hogrefe.
- Marx, P. (2007). *Lese- und Rechtschreiberwerb*. Paderborn: Schöningh.
- Marx, P. & Lenhard, W. (2011). Diagnostische Merkmale von Screening-Verfahren zur Früherkennung möglicher Probleme beim Schriftspracherwerb. In M. Hasselhorn & W. Schneider (Hrsg.), *Frühprognose schulischer Kompetenzen* (Jahrbuch der pädagogisch-psychologischen Diagnostik, Tests und Trends, Neue Folge Bd. 9, S. 68-84). Göttingen: Hogrefe.
- May, P. (2001). Diagnose der Rechtschreibstrategien mit der Hamburger Schreibprobe. In I. M. Naegele & R. Valtin (Hrsg.), *Schulische Förderung und außerschulische Therapien* (LRS – Legasthenie in den Klassen 1-10. Handbuch der Lese-Rechtschreibschwierigkeiten, Bd. 2, 2., überarbeitete Aufl., S. 87-92). Weinheim: Beltz.
- May, P. (unter Mitarbeit von Vieluf, U. & Malitzky, V.) (2002). *HSP 1-9. Diagnose orthographischer Kompetenz. Zur Erfassung der grundlegenden Rechtschreibstrategien mit der Hamburger Schreibprobe. Neustandardisierung 2001* (6., aktualisierte und erweiterte Aufl.). Hamburg: vpm.
- May, P. (2005a). *HSP 1+. Hamburger Schreib-Probe für die Klassenstufen 1/2 (Mitte Klasse 1, Ende Klasse 1, Mitte Klasse 2). Hinweise zur Durchführung und Auswertung. Neustandardisierung 2001*. Hamburg: vpm.
- May, P. (2005b). *HSP 2. Hamburger Schreib-Probe für die Klasse 2 (Ende Klasse 2). Hinweise zur Durchführung und Auswertung. Neustandardisierung 2001*. Hamburg: vpm.
- May, P. (2008). Diagnose der orthografischen Kompetenz – von der HSP zur DSP. In W. Schneider, H. Marx & M. Hasselhorn (Hrsg.), *Diagnostik von Rechtschreibleistungen und -kompetenz* (S. 93-127). Göttingen: Hogrefe.
- McLane, K. (n. d.). *Integrating Student Progress Monitoring into Your Classroom: The Teacher's Perspective* (Practitioner Brief No. 3). Zugriff am 08.09.2010. Verfügbar unter <http://www.studentprogress.org/doc/PractitionerBrief3FINAL.pdf>
- Menzel, W. (2003). *Pustelblume. Das Sprachbuch. 2. Schuljahr. Neubearbeitung*. Braunschweig: Schroedel.
- Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, Senatsverwaltung für Bildung, Jugend und Sport Berlin, Senator für Bildung und Wissenschaft Bremen & Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern. (2004). *Rahmenplan Grundschule Deutsch*. Zugriff am 17.06.2009. Verfügbar unter <http://www.bildungsserver-mv.de/download/rahmenplaene/rp-deutsch-gs.pdf>

- Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern. (2009a). *Schulbuchkatalog für allgemein bildende und berufliche Schulen in Mecklenburg-Vorpommern für das Schuljahr 2009/2010*. Zugriff am 16.06.2009. Verfügbar unter http://service.mvnet.de/_php/download.php?datei_id=9143
- Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern. (2009b). *Schulgesetz für das Land Mecklenburg-Vorpommern. Vom 13. Februar 2006 in der Fassung des ersten Änderungsgesetzes vom 16. Februar 2009*. Schwerin: Herausgeber.
- Moosbrugger, H. (2007). Item-Response-Theorie (IRT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 215-259). Heidelberg: Springer.
- Moosbrugger, H. & Schermelleh-Engel, K. (2007). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 307-324). Heidelberg: Springer.
- Moser Opitz, E. (2006). Assessments, Förderplanung, Förderdiagnostik – messen und/oder fördern? *Schweizerische Zeitschrift für Heilpädagogik*, 12 (9), 5-11.
- Müller, C. & Hartmann, E. (2009). Lernfortschritte im Unterricht erheben – Möglichkeiten und Grenzen des curriculumbasierten Messens. *Schweizerische Zeitschrift für Heilpädagogik*, 15 (10), 36-42.
- Müller, R. (1990). *DRT 1. Diagnostischer Rechtschreibtest für 1. Klassen*. Weinheim: Beltz.
- Müller, R. (1997). *DRT 2. Diagnostischer Rechtschreibtest für 2. Klassen. Neunormierung 1996* (3., neu normierte Aufl.). Weinheim: Beltz.
- Müller, R. (2004). *DRT 2. Diagnostischer Rechtschreibtest für 2. Klassen* (4., aktualisierte Aufl. in neuer Rechtschreibung). Göttingen: Beltz.
- National Association of State Directors of Special Education (n. d.). *Response to Intervention. Policy Considerations and Implementation*. Zugriff am 03.03.2009. Verfügbar unter http://www.nasdse.org/Portals/0/Documents/Download%20Publications/NASDSE_RtI.ppt
- National Center for Learning Disabilities (2006). *A Parent's Guide to Response-to-Intervention*. Zugriff am 03.09.2010. Verfügbar unter <http://www.ncld.org/images/stories/Publications/AdvocacyBriefs/ParentGuide-RTI/ParentsGuidetoRTI.pdf>
- National Joint Committee on Learning Disabilities (2005). *Responsiveness to Intervention and Learning Disabilities*. Zugriff am 16.04.2009. Verfügbar unter <http://www.ldanatl.org/pdf/rti2005.pdf>
- National Research Center on Learning Disabilities (2007). *Core Concepts of RTI*. Zugriff am 11.03.2009. Verfügbar unter <http://www.nrcld.org/about/research/rti/concepts.html>
- Naumann, C.L. (2008). Zur Rechtschreibkompetenz und ihrer Entwicklung. In A. Bremerich-Vos, D. Granzer & O. Köller (Hrsg.), *Lernstandsbestimmung im Fach*

- Deutsch. Gute Aufgaben für den Unterricht* (S. 134-159). Weinheim: Beltz.
- Neef, M. (2005a). Didaktische Konsequenzen einer zweistufigen Konzeption des Schriftsystems (am Beispiel von Schärfungsdaten). In H.-W. Huneke (Hrsg.), *Geschriebene Sprache. Strukturen, Erwerb, didaktische Modellbildungen* (S. 29-52). Heidelberg: Mattes.
- Neef, M. (2005b). *Die Graphematik des Deutschen*. Tübingen: Niemeyer.
- Neef, M. & Weingarten, R. (Hrsg.). (in Druck). *Schriftlinguistik. Ein Lern- und Konsultationswörterbuch mit systematischer Einleitung und englischen Übersetzungen* (Wörterbuch zur Sprach- und Kommunikationswissenschaft, Bd. 5). Berlin: de Gruyter.
- Nerius, D. (2000). *Beiträge zur deutschen Orthographie. Herausgegeben von Petra Ewald und Bernd Skibitzki anlässlich des 65. Geburtstages von Dieter Nerius*. Frankfurt am Main: Peter Lang.
- Nerius, D. (2007). *Deutsche Orthographie* (4., neu bearbeitete Aufl.). Hildesheim: Olms.
- Nickel, S. (2006). *Orthographieerwerb und die Entwicklung von Sprachbewusstheit. Zu Genese und Funktion von orthographischen Bewusstseinsprozessen beim frühen Rechtschreiberwerb in unterschiedlichen Lernkontexten*. Norderstedt: Books on Demand.
- Olson, A. & Caramazza, A. (1994). Representation and Connectionist Models: The NETspell Experience. In G. D. A. Brown & N. C. Ellis (Hrsg.), *Handbook of Spelling. Theory, Process and Intervention* (S. 337-363). Chichester: John Wiley & Sons.
- Parmar, R. S. (2008). Lernbehinderung – Ein Forschungsüberblick zur Situation in den USA am Beispiel von Lesen und Mathematik. *Vierteljahrszeitschrift für Heilpädagogik und ihre Nachbargebiete*, 77, 43-56.
- Petermann, F. (2000). Prozessevaluation und Verlaufsanalysen. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften. Themen, Strategien und Methoden der Rehabilitationsforschung* (S. 347-360). Berlin: Springer.
- Ramers, K.-H. (2001). *Einführung in die Phonologie* (2. Aufl.). München: Fink.
- Rapp, B., Epstein, C. & Tainturier, M.-J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology*, 19, 1-29.
- Rat für deutsche Rechtschreibung (2006). *Deutsche Rechtschreibung. Regeln und Wörterverzeichnis. Amtliche Regelung*. Tübingen: Narr Francke Attempto.
- Rauch, D. & Hartig, J. (2007). Interpretation von Testwerten in der IRT. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 240-250). Heidelberg: Springer.
- Reichen, J. (2007). Was ist „Lesen durch Schreiben“? Textsammlung Dr. Jürgen Reichen (ab 1981). Zugriff am 26.07.2012. Verfügbar unter <http://www.heinevetter->

- Reisinger, C.-M. (2007). *Unterrichtsdifferenzierung*. Berlin: Lit.
- Rentsch, K. & Schütz, A. (2009). *Psychologische Diagnostik. Grundlagen und Anwendungsperspektiven* (Grundriss der Psychologie, Bd. 16). Stuttgart: Kohlhammer.
- Reschly, D. J. (2005). Learning Disabilities Identification: Primary Intervention, Secondary Intervention, and Then What? *Journal of Learning Disabilities*, 38, 510-515.
- Richter, S. (2007). Lesen lernen, Schreiben lernen und Lese-Rechtschreibschwäche. In T. Fleischer, N. Grewe, B. Jötten, K. Seifried, B. Sieland (Hrsg.), *Handbuch Schulpsychologie. Psychologie für die Schule* (S. 151-160). Stuttgart: Kohlhammer.
- Rieder, O. (1991). *AST 2. Allgemeiner Schulleistungstest für 2. Klassen. Beiheft mit Anleitung und Normentabellen*. Weinheim: Beltz.
- Rost, J. (1996). *Lehrbuch Testtheorie Testkonstruktion*. Bern: Huber.
- Salomon, G. (1972). Heuristic Models for the Generation of Aptitude-Treatment Interaction Hypothesis. *Review of Educational Research*, 42, 327-343.
- Salomon, G. (1975). Heuristische Modelle für die Gewinnung von Interaktionshypothesen. In R. Schwarzer & K. Steinhagen (Hrsg.), *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (S. 127-145). München: Kösel.
- Salvia, J. & Ysseldyke, J. E. (unter Mitarbeit von Bolt, S.) (2007). *Assessment In Special and Inclusive Education* (10. Aufl.). Boston, MA: Houghton Mifflin Company.
- Saussure, F. de (2001). *Grundfragen der allgemeinen Sprachwissenschaft* (3. Aufl.). Berlin: de Gruyter.
- Schabmann, A. & Schmidt, B. M. (2009). Sind Lehrer gute Lese-Rechtschreibdiagnostiker? Der Einfluss von problematischem Schülerverhalten auf die Einschätzung der Lesekompetenz durch Lehrkräfte. *Heilpädagogische Forschung*, 35, 133-145.
- Scheerer-Neumann, G. (1986). Wortspezifisch: Ja – Wortbild: Nein. Ein letztes Lebewohl an die Wortbildtheorie. Teil I: Rechtschreiben. In H. Brügelmann (Hrsg.), *ABC und Schriftsprache: Rätsel für Kinder, Lehrer und Forscher* (S. 171-185). Konstanz: Faude.
- Scheerer-Neumann, G. (1987). Kognitive Prozesse beim Rechtschreiben: Eine Entwicklungsstudie. In G. Eberle & G. Reiß (Hrsg.), *Probleme beim Schriftspracherwerb. Möglichkeiten ihrer Vermeidung und Überwindung* (S. 193-219). Heidelberg: Edition Schindele.
- Scheerer-Neumann, G. (1996). Der Erwerb der basalen Lese- und Schreibfähigkeiten. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 10, Halbbd. 2, S. 1153-1169). Berlin: de Gruyter.

- Scheerer-Neumann, G. (1998). Schriftspracherwerb: „The State of the Art“ aus psychologischer Sicht. In L. Huber, G. Kegel & A. Speck-Hamdan (Hrsg.), *Einblicke in den Schriftspracherwerb* (S. 31-46). Braunschweig: Westermann.
- Scheerer-Neumann, G. (2003). Rechtschreibschwäche im Kontext der Entwicklung. In I. M. Naegele & R. Valtin (Hrsg.), *Grundlagen und Grundsätze der Lese-Rechtschreib-Förderung* (LRS – Legasthenie in den Klassen 1-10. Handbuch der Lese-Rechtschreibschwierigkeiten, Bd. 1, 6., vollständig überarbeitete Aufl., S. 45-65). Weinheim: Beltz.
- Scheerer-Neumann, G. (2006, 2. Quartal). Strukturen erkennen – Wörter üben. *Grundschule Deutsch*, (10), 22-25.
- Scheerer-Neumann, G. (2007). Rechtschreiben. In J. Walter & F. B. Wember (Hrsg.), *Sonderpädagogik des Lernens* (Handbuch Sonderpädagogik, Bd. 2, S. 539-568). Göttingen: Hogrefe.
- Scheerer-Neumann, G. (2008). Frühe Rechtschreibförderung zur Vorbeugung von Rechtschreibschwäche. In J. Borchert, B. Hartke & P. Jogschies (Hrsg.), *Frühe Förderung entwicklungsauffälliger Kinder und Jugendlicher* (S. 164-177). Stuttgart: Kohlhammer.
- Scheerer-Neumann, G., Schnitzler, C. D. & Ritter, C. (2009). *ILeA 2. Individuelle Lernstandsanalysen. Rechtschreiben. Lehrerheft*. Zugriff am 23.07.2010. Verfügbar unter http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/unterricht/lernstandsanalysen_vergleichsarbeiten/ilea/pdf/2LehrerSchreiben_Druck_.pdf
- Schermelleh-Engel, K. & Werner, C. (2007). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 113-133). Heidelberg: Springer.
- Schlee, J. (2004). Lösungsversuche als Problem. Zur Vergeblichkeit der so genannten Förderdiagnostik. In W. Mutzeck & P. Jogschies (Hrsg.), *Neue Entwicklungen in der Förderdiagnostik* (S. 23-38). Weinheim: Beltz.
- Schlee, J. (2007). „Förderdiagnostik ist ein verhängnisvoller Mythos“. Interview mit Jörg Schlee. *Heilpädagogik online*, 6 (2), 59-66. Verfügbar unter http://www.heilpaedagogik-online.com/2007/heilpaedagogik_online_0207.pdf
- Schlee, J. (2008). 30 Jahre »Förderdiagnostik« - eine kritische Bilanz. *Zeitschrift für Heilpädagogik*, 59, 122-130.
- Schneider, W. & Stefanek, J. (2007). Entwicklung der Rechtschreibleistung vom frühen Schul- bis zum frühen Erwachsenenalter. Längsschnittliche Befunde der Münchner LOGIK-Studie. *Zeitschrift für Pädagogische Psychologie*, 21, 77-82.
- Schneider, W., Stefanek, J. & Dotzler, H. (1997). Erwerb des Lesens und des Rechtschreibens: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 113-129). Weinheim: Beltz.

- Schöler, H. (2008, September). *Vom Unsinn der Förderdiagnostik. Vortrag anl. Kongress der Deutschen Gesellschaft für Sprachheilpädagogik (dgs), Cottbus*. Zugriff am 20.09.2010. Verfügbar unter http://www.ph-heidelberg.de/wp/schoeler/Datein/Cottbus-2008_Vom%20Unsinn%20der%20F%C3%B6rderdiagnostik_27-09-2008.pdf
- Schönweiss, F. (2008). *Lernserver. Interaktive Förderdiagnostik. Handbuch zur Rechtschreibförderung. Grundlagen und Förderpraxis* (2. Aufl.). Donauwörth: Auer.
- Schulz, W., Dertmann, J. & Jagla, A. (2003). Kinder mit Lese- und Rechtschreibstörungen: Selbstwertgefühl und Integrative Lerntherapie. *Kindheit und Entwicklung*, 12, 231-242.
- Schwarzer, G., Türp, J. & Antes, G. (2002). Wahrscheinlichkeitsverhältnis (Likelihood Ratio) - Alternative zu Sensitivität und Spezifität. *Deutsche Zahnärztliche Zeitschrift*, 57, 660-661.
- Schwarzer, R. & Steinhagen, K. (1975). Adaptiver Unterricht als Beitrag zu einer pädagogischen Ökologie. In R. Schwarzer & K. Steinhagen (Hrsg.), *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (S. 11-26). München: Kösel.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. Zugriff am 30.07.2010. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Deutsch-Primar.pdf
- Senate and House of Representatives of the United States of America in Congress (2002, January 8). *No Child Left Behind Act of 2001* (Public Law 107-110, 107th Congress). Zugriff am 02.03.2009. Verfügbar unter <http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>
- Senate and House of Representatives of the United States of America in Congress (2004, December 3). *Individuals with Disabilities Education Improvement Act of 2004* (Public Law 108-446, 108th Congress). Zugriff am 02.03.2009. Verfügbar unter <http://ies.ed.gov/ncser/pdf/pl108-446.pdf>
- Shapiro, E. S. (2009). *The Two Models of RTI: Standard Protocol and Problem Solving*. Virginia: Department of Education, Office of Special Education and Student Services. Zugriff am 07.09.2010. Verfügbar unter http://www.doe.virginia.gov/instruction/response_intervention/guidance/two_models.pdf
- Shinn, M. R. (1981). *A comparison of psychometric and functional differences between students labeled learning disabled and low achieving*. Unpublished doctoral dissertation, University of Minneapolis, MN.
- Shinn, M. R. & Shinn, M. M. (2002). *AIMSweb® Training Workbook: Administration and Scoring of Spelling Curriculum-Based Measurement (S-CBM) for Use in General Outcome Measurement*. Eden Prairie, MN: Edformation. Zugriff am 22.09.2010. Verfügbar unter http://www.aimsweb.com/uploads/pdfs/scoring_scbm.pdf
- Shores, C. & Bender, W. N. (2007). Response to Intervention. In W. N. Bender & C.

- Shores (Hrsg.), *Response to Intervention. A Practical Guide for Every Teacher* (S. 1-19). Thousand Oaks, CA: Corwin Press.
- Simon, D. P. (1976). Spelling – A task analysis. *Instructional Science*, 5, 277-302.
- Simon, D. P. & Simon, H. A. (1973). Alternative Uses of Phonemic Information in Spelling. *Review of Educational Research*, 43, 115-137.
- Snow, R. E. (1980). Aptitude Processes. In R. E. Snow, P.-A. Federico & W. E. Montague (Hrsg.), *Aptitude, Learning and Instruction. Cognitive Process Analyses of Aptitude* (Bd. 1, S. 27-63). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Snow, R. E. (1982). Education and intelligence. In R. J. Sternberg (Hrsg.), *Handbook of human intelligence* (S. 493-585). New York, NY: Cambridge University Press.
- Snow, R. E. (1989). Aptitude-Treatment Interaction as a Framework for Research on Individual Differences in Learning. In P. L. Ackerman, R. J. Sternberg & R. Glaser (Hrsg.), *Learning and individual differences. Advances in theory and research* (S. 13-59). New York, NY: W.H. Freeman and Company.
- Snow, R. E. (1991a). Aptitude-Treatment Interaction as a Framework for Research on Individual Differences in Psychotherapy. *Journal of Consulting and Clinical Psychology*, 59, 205-216.
- Snow, R. E. (1991b). The Concept of Aptitude. In R. E. Snow & D. E. Wiley (Hrsg.), *Improving inquiry in social science. A Volume in Honor of Lee J. Cronbach* (S. 249-284). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Souvignier, E. & Förster, N. (2011). Effekte prozessorientierter Diagnostik auf die Entwicklung der Lesekompetenz leseschwacher Viertklässler. *Empirische Sonderpädagogik*, 3, 243-255.
- Speece, D. L., Case, L. P. & Molloy, D. E. (2003). Responsiveness to General Education Instruction as the First Gate to Learning Disabilities Identification. *Learning Disabilities Research & Practice*, 18, 147-156.
- Stock, C. & Schneider, W. (2008a). *DERET 1-2+. Deutscher Rechtschreibtest für das erste und zweite Schuljahr*. Göttingen: Hogrefe.
- Stock, C. & Schneider, W. (2008b). Die Deutschen Rechtschreibtests für das Grundschulalter (DERET 1-2+ und DERET 3-4+). In W. Schneider, H. Marx & M. Hasselhorn (Hrsg.), *Diagnostik von Rechtschreibleistungen und -kompetenz* (S. 45-60). Göttingen: Hogrefe.
- Strathmann, A. M. & Klauer, K. J. (2008). Diagnostik des Lernverlaufs. Eine Pilotstudie am Beispiel der Entwicklung der Rechtschreibkompetenz. *Sonderpädagogik*, 38, 5-24.
- Strathmann, A. M. & Klauer, K. J. (2010). Lernverlaufsdiagnostik: Ein Ansatz zur längerfristigen Lernfortschrittsmessung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42, 111-122.

- Strathmann, A. M. & Klauer, K. J. (in Vorbereitung). *Lernverlaufsdiagnostik Mathematik*. Göttingen: Hogrefe.
- Strathmann, A. M., Klauer, K. J. & Greisbach, M. (2010). Lernverlaufsdiagnostik – Dargestellt am Beispiel der Entwicklung der Rechtschreibkompetenz in der Grundschule. *Empirische Sonderpädagogik*, 2, 64-77.
- Szelenko, C., Schmidt, T. & Wessel, H. (2004). *Sprachfreunde 2. Ein Sprachbuch für die Grundschule*. Berlin. Volk und Wissen.
- Thomé, G. (2006). Entwicklung der basalen Rechtschreibkenntnisse. In U. Bredel, H. Günther, P. Klotz, J. Ossner & G. Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache* (Bd. 1, 2., durchgesehene Aufl., S. 369-379). Paderborn: Schöningh.
- Thomé, G. & Thomé, D. (2004). *OLFA. Oldenburger Fehleranalyse. Instrument und Handbuch zur Ermittlung der orthographischen Kompetenz aus freien Texten ab Klasse 3 und zur Qualitätssicherung von Fördermaßnahmen*. Oldenburg: Igel Verlag Wissenschaft.
- Tindal, G., Germann, G. & Deno, S. L. (1983). *Descriptive research on the Pine County norms: A compilation of findings* (Research Report No. 132). Minneapolis, MN: University, Institute for Research on Learning Disabilities.
- Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology*, 83, 346-360.
- U.S. Department of Education (2002, July 1). *A New Era: Revitalizing Special Education for Children and Their Families. President's Commission on Excellence in Special Education*. Zugriff am 04.03.2009. Verfügbar unter http://www.ed.gov/inits/commissionsboards/whspecialeducation/reports/images/Pres_Rep.pdf
- U.S. Department of Education (2006, August 14). 34 CFR Parts 300 and 301. RIN 1820–AB57. Assistance to States for the Education of Children With Disabilities and Preschool Grants for Children With Disabilities. *Federal Register*, 71, 465539-46845. Zugriff am 02.03.2009. Verfügbar unter <http://idea.ed.gov/download/finalregulations.pdf>
- Valtin, R., Meyer-Schepers, U. & Löffler, I. (2003). Rechtschreiben – ein wahres Schulkreuz? *Grundschule*, 35 (12), 40-41.
- Valtin, R. & Wagner, C. (2002). Wie wirken sich Notengebung und verbale Beurteilung auf die leistungsbezogene Persönlichkeitsentwicklung aus? In R. Valtin (Hrsg.), *Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand* (S. 113-137). Weinheim: Juventa.
- Vaughn, S. & Fuchs, L. S. (2003). Redefining Learning Disabilities as Inadequate Response to Instruction: The Promise and Potential Problems. *Learning Disabilities Research & Practice*, 18, 137-146.
- Vaughn, S., Wanzek, J., Woodruff, A. L. & Linan-Thompson, S. (2007). Prevention and

- Early Identification of Students with Reading Disabilities. In D. Haager, J. Klingner & S. Vaughn (Hrsg.), *Evidence-Based Reading Practices for Response to Intervention* (S. 11-27). Baltimore, MD: Paul. H. Brookes Publishing Co.
- Voss, A., Löffler, I., Meyer-Schepers, U., Meckel, C. & Kowalski, K. (2008). Frühdiagnose rechtschreibschwächerer Schülerinnen und Schüler auf der Grundlage von Kompetenzmodellen. Die Analyse von Lernentwicklungsverläufen als Aufgabe schulischer Effektivitätsforschung. In W. Bos, H. G. Holtappels, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung* (Bd. 15, S. 123-155). Weinheim: Juventa.
- Walter, J. (2008a). Adaptiver Unterricht erneut betrachtet: Über die Notwendigkeit systematischer formativer Evaluation von Lehr- und Lernprozessen und die daraus resultierende Diagnostik und Neudefinition von Lernstörungen nach dem RTI-Paradigma. *Zeitschrift für Heilpädagogik*, 59, 202-215.
- Walter, J. (2008b). Curriculumbasiertes Messen (CBM) als lernprozessbegleitende Diagnostik: Erste deutschsprachige Ergebnisse zur Validität, Reliabilität und Veränderungssensibilität eines robusten Indikators zur Lernfortschrittsmessung beim Lesen. *Heilpädagogische Forschung*, 34, 62-79.
- Walter, J. (2009a). Eignet sich die Messtechnik „MAZE“ zur Erfassung von Lesekompetenzen als lernprozessbegleitende Diagnostik? *Heilpädagogische Forschung*, 35, 62-75.
- Walter, J. (2009b). Theorie und Praxis Curriculumbasierten Messens (CBM) in Unterricht und Förderung. *Zeitschrift für Heilpädagogik*, 9, 162-170.
- Walter, J. (2010). *Lernfortschrittsdiagnostik Lesen (LDL). Ein curriculumbasiertes Verfahren*. Göttingen: Hogrefe.
- Weingarten, R. (2003). Schriftspracherwerb. In G. Rickheit, T. Herrmann & W. Deutsch (Hrsg.), *Psycholinguistik. Ein internationales Handbuch* (Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 24, S. 801-811). Berlin: de Gruyter.
- Wember, F. B. (2001). Adaptiver Unterricht. *Sonderpädagogik*, 31, 161-181.
- Westfälische Wilhelms-Universität Münster. (n.d.). *Testübersicht. Quop. Plattform für Schule und Entwicklung*. Zugriff am 25.02.2012. Verfügbar unter http://www.quop.de/fileadmin/uploads/quop/literature/quop_Test%C3%BCbersicht.pdf
- Wilbert, J. & Linnemann, M. (2011). Kriterien zur Analyse eines Tests zur Lernverlaufsdiagnostik. *Empirische Sonderpädagogik*, 3, 225-242.
- Woolfolk, A. (2008). *Pädagogische Psychologie* (10. Aufl.). München: Pearson Studium.
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3, 32-35. Zugriff am 10.10.2011. Verfügbar unter [http://onlinelibrary.wiley.com/doi/10.1002/1097-0142\(1950\)3:1%3C32::AID-CNCR2820030106%3E3.0.CO;2-3/pdf](http://onlinelibrary.wiley.com/doi/10.1002/1097-0142(1950)3:1%3C32::AID-CNCR2820030106%3E3.0.CO;2-3/pdf)
- Zendler, A., Spannagel, C. & Vogel, M. (2008). Versuchspläne mit Messwiederholung

unter Verwendung kleiner Stichproben - Teil 1: Schülermerkmale. *Notes on Educational Informatics - Section A: Concepts and Techniques*, 4 (1), 15-25. Zugriff am 17.02.2011. Verfügbar unter http://www.ph-ludwigsburg.de/fileadmin/subsites/2e-imix-t-01/user_files/Journal_NEI_-_PDFs_fuer_Webauftritt/Section_A/Volume_4_No_1_2008/NEI_Section_A_Vol_4_No_1_2008_p_15-25_-_Zendler_-_Versuchsplaene_mit_Messwiederholung_Teil_I.pdf

Abkürzungsverzeichnis

A	Aptitude-Treatment Interaction
A ₁ /A ₂	Aptitude
A ₁	Form A zum 1. Messzeitpunkt
A ₂	Form A zum 2. Messzeitpunkt
A ₃	Form A zum 3. Messzeitpunkt
A ₄	Form A zum 4. Messzeitpunkt
AFRA	Aachener Förderdiagnostische Rechtschreibfehler-Analyse
AST 2	Allgemeine Schulleistungstest für zweite Klassen
B	unstandardisiertes Regressionsgewicht
B ₁	Form B zum 1. Messzeitpunkt
B ₂	Form B zum 2. Messzeitpunkt
B ₃	Form B zum 3. Messzeitpunkt
B ₄	Form B zum 4. Messzeitpunkt
β	standardisiertes Regressionsgewicht
B ₁ /B ₂	Treatment
CBA	Curriculum-Based Assessment
CBE	Curriculum-Based Evaluations
CBM	Curriculum-Based Measurement
CFT 1	Grundintelligenz Skala 1
CI	Konditionsindex
d	Durbin-Watson-Koeffizient
DEMAT 1+	Deutscher Mathematiktest für erste Klassen
DEMAT 2+	Deutscher Mathematiktest für zweite Klassen
DERET 1-2+	Deutscher Rechtschreibtest für das erste und zweite Schuljahr
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
DoRA	Dortmunder Rechtschreibfehler Analyse
DRT 1	Diagnostischer Rechtschreibtest für 1. Klassen
DRT 2	Diagnostischer Rechtschreibtest für 2. Klassen
EFA	explorative Faktorenanalyse
E _g	graphematische Ebene
ELG	elementar graphematisches Können, Lupenstelle der elementar graphematischen Stufe

ELO	elementar orthographisches Können, Lupenstelle der elementar orthographischen Stufe
E _p	phonologische Ebene
ERG	erweitert graphematisches Können, Lupenstelle der erweitert graphematischen Stufe
ERO	erweitert orthographisches Können, Lupenstelle der erweitert orthographischen Stufe
E _s	semantische Ebene
et al.	et alii
FE-AF 2	Formative Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr
FE-L 2	Formative Erfassung der Lesefertigkeit im 2. Schuljahr
FE-RS 2	Formative Erfassung der Rechtschreibleistung im 2. Schuljahr
Form A	Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A
Form B	Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B
Fortg.	Fortgeschrittener
GOM	General Outcome Measurement
GS	Grundschule
H	Hypothese
HSP 1+	Hamburger Schreib-Probe für die Klassenstufen 1/2
HSP 2	Hamburger Schreib-Probe für die Klasse 2
IDEIA	Individuals with Disabilities Education Improvement Act of 2004
K	Konsonant
KFA	konfirmative Faktorenanalyse
KMK	Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland
KMO	Kaiser-Meyer-Olkin-Koeffizient
KST	Kolmogorov-Smirnov-Test
KT	Klassische Testtheorie
LR	Likelihood-Ratio
LR+	positiver Likelihood-Ratio
LR-	negativer Likelihood-Ratio
LD	learning disabilities
LOVS	Leerlingonderwijsvolgsysteem
M	Mittelwert

männl.	Männlich
Max.	Maximum
MBJS	Ministerium für Bildung, Jugend und Sport des Landes Brandenburg
MBWK	Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern
Md	Median
Min.	Minimum
MM	Mastery Measurement
MRA	Münsteraner Rechtschreibanalyse
MSA	Measure of Sample Adequacy-Koeffizient
MST	Mecklenburg-Strelitz
MZP	Messzeitpunkt
N	Anzahl
NB	Neubrandenburg
NCLB	No Child Left Behind Act of 2001
n. d.	nicht datiert
NPW	negativer prädiktiver Wert in Prozent
Nr.	Ordnungsnummer des Wortmaterials
NRC	National Research Council
OG	orthographischer Gesamtwert
OLFA	Oldenburger Fehleranalyse
p	Irrtumswahrscheinlichkeit
(P)	Korrelationskoeffizient nach Pearson
PGK	Phonem-Graphem-Korrespondenz
p_i	Itemschwierigkeit
PPW	positiver prädiktiver Wert in Prozent
PR	Prozentrang
PTT	Probabilistische Testtheorie
r	Korrelationskoeffizient
R^2_{korr}	korrigiertes R^2 = Bestimmtheitsmaß
RATZ	Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote
R-I	Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Prozent
r_{it}	Trennschärfekoeffizient

RSG	Regionale Schule mit Grundschulteil
RS-Note	Rechtschreibnote
RTI	Response to Intervention
r_{tt}	Reliabilitätskoeffizient
s.	signifikanter Korrelationsunterschied
(S)	Korrelationskoeffizient nach Spearman
SBJS	Senatsverwaltung für Bildung, Jugend und Sport Berlin
SBL II	Schultestbatterie zur Erfassung des Lernstandes in Mathematik, Lesen und Schreiben II
SBW	Senator für Bildung und Wissenschaft Bremen
SD	Standardabweichung
SE	Standardmessfehler
SEN	Sensitivität in Prozent
SLRT	Salzburger Lese- und Rechtschreibtest
SPE	Spezifität in Prozent
V	Vokal
weibl.	Weiblich
WLLP	Würzburger Leiseleseprobe
WRT 1+	Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen
WRT 2+	Weingartener Grundwortschatz Rechtschreib-Test für zweite und dritte Klassen
y	Outcome
Y-I	Youden-Index
z	Prüfgröße
Z	zentrales Phonem
z_{krit}	kritischer Grenzwert von z

Tabellenverzeichnis

Tabelle 1:	Adaptionsmaßnahmen (in Anlehnung an Leutner, 1992, S. 10).....	24
Tabelle 2:	Entwicklungsstufen beim Erwerb der Rechtschreibung (Scheerer-Neumann, 2008, S. 166)	75
Tabelle 3:	Entwicklungsmodell des graphematischen und orthographischen Rechtschreibkönnens (in Anlehnung an Löffler & Meyer-Schepers, 2008; May, 2002; Neef, 2005a, b; Scheerer-Neumann, 2006, 2008).....	83
Tabelle 4:	Ausgewählte Ziele und Inhalte aus dem Aufgabenbereich „Schreiben: Texte verfassen/Rechtschreiben“ für die Klassenstufe 1/2 (MBJS et al., 2004, S. 28).....	104
Tabelle 5:	Anforderungen und Inhalte des Lernbereichs Rechtschreiben in der 2. Klassenstufe	111
Tabelle 6:	Untersuchungsdesign	114
Tabelle 7:	Wortmaterial und Lupenstellen der FE-RS 2.....	121
Tabelle 8:	Erprobungsstichprobe der FE-RS 2.....	124
Tabelle 9:	Kennwerte des orthographischen Gesamtwerts der Form A in Klasse 2 und 3 im Rahmen der Erprobungsuntersuchung	125
Tabelle 10:	Kennwerte des orthographischen Gesamtwerts der Form B in Klasse 2 und 3 im Rahmen der Erprobungsuntersuchung	127
Tabelle 11:	Kriterien für signifikante und nicht signifikante Ergebnisse (Bühl, 2008, S. 121)	139
Tabelle 12:	Einteilung der Größen der Korrelationskoeffizienten (Bühl, 2008, S. 346)	140
Tabelle 13:	Wertebereiche der Wahrscheinlichkeitsverhältnisse und ihre Interpretation (G. Schwarzer, Türp & Antes, 2002, S. 660)	147
Tabelle 14:	Verteilung der Stichprobe nach Region und Schultyp	148
Tabelle 15:	Verteilung der Teilstichproben nach Region, Schultyp und Geschlecht	149
Tabelle 16:	Altersverteilung der Stichprobe zu den 4 Messzeitpunkten.....	149
Tabelle 17:	Kennwerte des orthographischen Gesamtwerts der Form A über alle 4 Messzeitpunkte.....	152

Tabelle 18:	Kennwerte der ELG-Subskala der Form A über alle 4 Messzeitpunkte	154
Tabelle 19:	Kennwerte der ELO-Subskala der Form A über alle 4 Messzeitpunkte	156
Tabelle 20:	Kennwerte der ERG-Subskala der Form A über alle 4 Messzeitpunkte	158
Tabelle 21:	Kennwerte des orthographischen Gesamtwerts der Form B über alle 4 Messzeitpunkte.....	160
Tabelle 22:	Kennwerte der ELG-Subskala der Form B über alle 4 Messzeitpunkte	162
Tabelle 23:	Kennwerte der ELO-Subskala der Form B über alle 4 Messzeitpunkte	164
Tabelle 24:	Kennwerte der ERG-Subskala der Form B über alle 4 Messzeitpunkte	166
Tabelle 25:	Gemittelte Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form A über alle 4 Messzeitpunkte.....	168
Tabelle 26:	Gemittelte Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form B über alle 4 Messzeitpunkte.....	169
Tabelle 27:	Retestreliabilität der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte	171
Tabelle 28:	Testhalbierungsreliabilität der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte	172
Tabelle 29:	Interne Konsistenz der FE-RS 2 für den orthographischen Gesamtwert und die Subskalen über alle 4 Messzeitpunkte	173
Tabelle 30:	Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2, dem DERET 1-2+ und der FE-L 2 zum 1. und 4. Messzeitpunkt	175
Tabelle 31:	Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2, dem DEMAT 1+, dem DEMAT 2+ und der FE-AF 2 zum 1. und 4. Messzeitpunkt	177
Tabelle 32:	Korrelationen zwischen dem orthographischen Gesamtwert der FE-RS 2 und der Rechtschreibnote zum 4. Messzeitpunkt	185
Tabelle 33:	Korrelationen zwischen den orthographischen Gesamtwerten	

	der FE-RS 2	186
Tabelle 34:	Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form A zum 3. Messzeitpunkt durch den orthographischen Gesamtwert der Form A zum 1. Messzeitpunkt (N = 193).....	187
Tabelle 35:	Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form A zum 4. Messzeitpunkt durch den orthographischen Gesamtwert der Form A zum 2. Messzeitpunkt (N = 232).....	188
Tabelle 36:	Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form B zum 3. Messzeitpunkt durch den orthographischen Gesamtwert der Form B zum 1. Messzeitpunkt (N = 193).....	188
Tabelle 37:	Regressionsanalyse zur Vorhersage des orthographischen Gesamtwerts der Form B zum 4. Messzeitpunkt durch den orthographischen Gesamtwert der Form B zum 2. Messzeitpunkt (N = 232).....	189
Tabelle 38:	Anzahl der Risikoschüler der FE-RS 2 zum 1. und 2. Messzeitpunkt aufgrund kritischer Prädiktoren-Testwerte (PR = 25, $N_{A1/B2} = 193$, $N_{B1/A2} = 232$).....	190
Tabelle 39:	Anzahl der Problemschüler der FE-RS 2 zum 3. und 4. Messzeitpunkt aufgrund kritischer Kriterien-Testwerte (PR = 15 für OG, RS-Note ≥ 4 , PR = 16 für Anzahl der Rechtschreibfehler im DERET 1-2+, $N_{A3/B4} = 193$, $N_{B3/A4} = 232$)	191
Tabelle 40:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 1. Messzeitpunkt und des Kriteriums zum 3. Messzeitpunkt.....	192
Tabelle 41:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 2. Messzeitpunkt und des Kriteriums zum 4. Messzeitpunkt.....	192
Tabelle 42:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt.....	193
Tabelle 43:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt.....	193
Tabelle 44:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt.....	194

Tabelle 45:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium	195
Tabelle 46:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium	195
Tabelle 47:	Kennwerte zur prognostischen Validität der Form A aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium	196
Tabelle 48:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 1. Messzeitpunkt und des Kriteriums zum 3. Messzeitpunkt	197
Tabelle 49:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 2. Messzeitpunkt und des Kriteriums zum 4. Messzeitpunkt	197
Tabelle 50:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt.....	198
Tabelle 51:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt.....	199
Tabelle 52:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Rechtschreibnote zum 4. Messzeitpunkt.....	199
Tabelle 53:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 1. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium	200
Tabelle 54:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 2. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium	201
Tabelle 55:	Kennwerte zur prognostischen Validität der Form B aufgrund der Prädiktoren zum 3. Messzeitpunkt und der Anzahl der Rechtschreibfehler im DERET 1-2+ zum 4. Messzeitpunkt als Kriterium	201
Tabelle 56:	Äquivalenzüberprüfung von Varianzen und Mittelwerten der	

	FE-RS 2	204
Tabelle 57:	Paralleltestreliabilität der FE-RS 2 zum 1. Messzeitpunkt (N = 68) und 4. Messzeitpunkt (N = 63)	205
Tabelle 58:	Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Retest-Koeffizienten der Form A und B.....	207
Tabelle 59:	Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Testhalbierungs-Koeffizienten der Form A und B	208
Tabelle 60:	Prüfung der Signifikanz von Korrelationsunterschieden zwischen den internen Konsistenzen der Form A und B	209
Tabelle 61:	Prüfung der Signifikanz von Korrelationsunterschieden zwischen den Validitätskoeffizienten der Form A und B	210
Tabelle 62:	Mittelwertvergleich des orthographischen Gesamtwerts und der Subskalen der Form A.....	214
Tabelle 63:	Kennwerte der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A.....	215
Tabelle 64:	Mittelwertvergleich des orthographischen Gesamtwerts der Form B.....	218
Tabelle 65:	Kennwerte der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form B	219
Tabelle 66:	Häufigkeiten der Antworten zur Frage „Nach welcher Methode unterrichten Sie im Lernbereich Rechtschreiben?“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)	220
Tabelle 67:	Häufigkeiten der Antworten zur Frage „Welche Sozialformen verwenden Sie im Lernbereich Rechtschreiben?“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)	221
Tabelle 68:	Häufigkeiten der Antworten zur Frage „Wie bewerten Sie die FE-RS 2? Bitte äußern Sie sich zu folgenden Aussagen!“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern)	221
Tabelle 69:	Häufigkeiten der Antworten zur Frage „Wie bewerten die Schüler Ihrer Meinung nach die FE-RS 2? Bitte äußern Sie sich zu folgenden Aussagen!“ (Angabe absoluter Häufigkeiten sowie relativer Häufigkeiten in Klammern).....	222
Tabelle 70:	Vergleich der FE-RS 2 mit in der deutschen Testzentrale	

aufgeführten Rechtschreibtests und Mehrfächertests mit einem Rechtschreibteil für die 2. Klassenstufe.....	236
---	-----

Abbildungsverzeichnis

Abbildung 1:	Die Interaktionsarten (in Anlehnung an R. Schwarzer & Steinhagen, 1975, S. 16).....	19
Abbildung 2:	RTI-Modell (Hartmann & C. Müller, 2009, S. 27).....	36
Abbildung 3:	A Progress-Monitoring Chart with an Aimline (Salvia & Ysseldyke, 2007, S. 648).....	38
Abbildung 4:	Beispiel für ein Rechtschreib-CBM der 3. Klassenstufe (AIMSweb, 2003).....	44
Abbildung 5:	Der Aufbau des Sprachsystems.....	54
Abbildung 6:	Der Silbenaufbau am Beispiel <i>Bein</i> (in Anlehnung an Grassegger, 2010).....	56
Abbildung 7:	The Six-step Model of Skills in Reading and Writing Acquisition (Frith, 1985, S. 311).....	65
Abbildung 8:	Entwicklung des Rechtschreibkönnens – Integration der grundlegenden Rechtschreibstrategien (May, 2002, S. 148)	71
Abbildung 9:	Kompetenzentwicklungsmodell des Rechtschreibens (in Anlehnung an Lischeid, 2007)	79
Abbildung 10:	Spelling by means of a generator-test process (D. P. Simon & H. A. Simon, 1973, S. 118)	86
Abbildung 11:	Routes to the orthographic output lexicon (Barry, 1994, S. 45).....	89
Abbildung 12:	Zwei-Komponenten-Modell der Rechtschreibung (Scheerer-Neumann, 2003, S. 51).....	92
Abbildung 13:	The architecture of the connectionist model. Only some units and connections are illustrated. (Brown & Loosemore, 1994, S. 323)	96
Abbildung 14:	Rechtschreibprozessmodell NETspell (in Anlehnung an Klicpera et al., 2007; Olson & Caramazza, 1994)	98
Abbildung 15:	Untersuchungsplan	116
Abbildung 16:	Häufigkeitsverteilung für den orthographischen Gesamtwert der Form A in Klasse 2 (N = 22) und in Klasse 3 (N = 23) im Rahmen der Erprobungsuntersuchung	124
Abbildung 17:	Häufigkeitsverteilung für den orthographischen Gesamtwert	

	der Form B in Klasse 2 (N = 21) und in Klasse 3 (N = 22) im Rahmen der Erprobungsuntersuchung	126
Abbildung 18:	Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form A	150
Abbildung 19:	Häufigkeitsverteilungen für die ELG-Subskala der Form A	152
Abbildung 20:	Häufigkeitsverteilungen für die ELO-Subskala der Form A	154
Abbildung 21:	Häufigkeitsverteilungen für die ERG-Subskala der Form A	156
Abbildung 22:	Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form B	159
Abbildung 23:	Häufigkeitsverteilungen für die ELG-Subskala der Form B	161
Abbildung 24:	Häufigkeitsverteilungen für die ELO-Subskala der Form B	163
Abbildung 25:	Häufigkeitsverteilungen für die ERG-Subskala der Form B	165
Abbildung 26:	Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	211
Abbildung 27:	Häufigkeitsverteilungen für die ELG-Subskala der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	212
Abbildung 28:	Häufigkeitsverteilungen für die ELO-Subskala der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	212
Abbildung 29:	Häufigkeitsverteilungen für die ERG-Subskala der Form A zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	213
Abbildung 30:	Häufigkeitsverteilungen für den orthographischen Gesamtwert der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	216
Abbildung 31:	Häufigkeitsverteilungen für die ELG-Subskala der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	216
Abbildung 32:	Häufigkeitsverteilungen für die ELO-Subskala der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	217
Abbildung 33:	Häufigkeitsverteilungen für die ERG-Subskala der Form B zwischen dem 1. und 3. sowie dem 2. und 4. Messzeitpunkt	217

Eidesstattliche Versicherung

Hiermit versichere an Eides statt, dass ich die eingereichte Dissertation mit dem Titel „Entwicklung und Evaluation eines Verfahrens zur Lernverlaufsdokumentation im Lernbereich Rechtschreiben der 2. Klassenstufe“ selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Ich versichere weiterhin, dass die vorliegende Dissertation weder insgesamt noch ausschnittsweise für die Erfüllung einer Auflage im Sinne von § 6, Absatz 2 und 5 der Promotionsordnung der Philosophischen Fakultät der Universität Rostock verwendet wurde und dass sie in keiner anderen akademischen oder staatlichen Prüfung vorgelegt wurde (§ 9, Absatz 7).

Rostock, den 02.01.2013

**Entwicklung und Evaluation eines Verfahrens zur
Lernverlaufsdokumentation im Lernbereich Rechtschreiben
der 2. Klassenstufe**

(Anhang)

**Dissertation
zur
Erlangung des akademischen Grades
Doctor philosophiae (Dr. phil.)
der Philosophischen Fakultät
der Universität Rostock**

vorgelegt von

**Kristin Kuhlmann, geb. am 31.10.1983 in Neubrandenburg
aus Köln**

Rostock, 02.01.2013

Anhangsverzeichnis

Anhang A:	Lehrerheft – Durchführungshinweise für die Formative Erfassung der Rechtschreibleistung (FE-RS 2), der Lesefertigkeit (FE-L 2) und der arithmetischen Fähigkeiten (FE-AF 2) im 2. Schuljahr.....	281
Anhang B:	Testheft der FE-RS 2 für die Schüler	290
Anhang C:	Lehrerfragebogen zur Einschätzung der diagnostischen Verfahren zur Lernfortschrittsmessung.....	292
Anhang D:	Lehrerfragebogen zur Einschätzung der Rechtschreibleistung der Schüler.....	294
Anhang E:	Vordruck der E-Mail mit Bitte um Unterstützung der Untersuchung	295
Anhang F:	Antrag auf Genehmigung der wissenschaftlichen Untersuchung an Grundschulen	296
Anhang G:	Mitteilung über die Antragsstellung.....	298
Anhang H:	Bewilligung der wissenschaftlichen Untersuchung durch das Schulamt Neubrandenburg	300
Anhang I:	Vordruck der Elterneinverständniserklärung	301
Anhang J:	Überprüfung auf Normalverteilung des orthographischen Gesamtwerts und der Subskalen der Form A mittels Kolmogorov-Smirnov-Test	302
Anhang K:	Überprüfung auf Normalverteilung des orthographischen Gesamtwerts und der Subskalen der Form B mittels Kolmogorov-Smirnov-Test	303
Anhang L:	Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form A über alle 4 Messzeitpunkte.....	304
Anhang M:	Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form B über alle 4 Messzeitpunkte.....	308
Anhang N:	Measure of Sample Adequacy-Koeffizienten für die Items der Form A über alle 4 Messzeitpunkte	312
Anhang O:	Hauptkomponentenanalysen der Form A zum 1. Messzeitpunkt.....	313

Anhang P:	Hauptkomponentenanalysen der Form A zum 2. Messzeitpunkt	319
Anhang Q:	Hauptkomponentenanalysen der Form A zum 3. Messzeitpunkt	325
Anhang R:	Hauptkomponentenanalysen der Form A zum 4. Messzeitpunkt	331
Anhang S:	Measure of Sample Adequacy-Koeffizienten für die Items der Form B über alle 4 Messzeitpunkte	337
Anhang T:	Hauptkomponentenanalysen der Form B zum 1. Messzeitpunkt	338
Anhang U:	Hauptkomponentenanalysen der Form B zum 2. Messzeitpunkt	344
Anhang V:	Hauptkomponentenanalysen der Form B zum 3. Messzeitpunkt	350
Anhang W:	Hauptkomponentenanalysen der Form B zum 4. Messzeitpunkt	356
Anhang X:	Regressionsanalysen der FE-RS 2	362
Anhang Y:	Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A.....	365
Anhang Z:	Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form B.....	368

**Anhang A: Lehrerheft – Durchführungshinweise für die Formative Erfassung
der Rechtschreibleistung (FE-RS 2), der Lesefertigkeit (FE-L 2) und
der arithmetischen Fähigkeiten (FE-AF 2) im 2. Schuljahr**

Formative Erfassung der Rechtschreibleistung (FE-RS 2), der
Leseferigkeit (FE-L 2) und der arithmetischen Fähigkeiten (FE-AF 2)
im 2. Schuljahr

Universität Rostock Hinrichsen/ Hartke

Lehrerheft
Durchführungshinweise

Universität Rostock
Institut für sonderpädagogische Entwicklungsförderung
und Rehabilitation
August-Bebel-Str. 28
18055 Rostock
Telefon: 0381-498 2678
sonderpaedagogik@uni-rostock.de

Kristin Hinrichsen
Neusser Wall 36
50668 Köln
Telefon: 0221/29818640
015775336426
kristinhinrichsen@gmx.de

Liebe Lehrerinnen und Lehrer,

ich freue mich sehr, dass Sie sich bereit erklären, die diagnostischen Verfahren zur formativen Erfassung der Rechtschreibleistung (FE-RS 2), der Lesefertigkeit (FE-L 2) und der arithmetischen Fähigkeiten (FE-AF 2) im 2. Schuljahr in Ihrer Klasse durchzuführen!

Die FE-RS 2, FE-L 2 und FE-AF 2 sind Diagnoseverfahren zur Lernfortschrittsmessung im Lernbereich Rechtschreiben, Lesen und Arithmetik der 2. Klasse in der Grundschule. Sie basieren auf aktuellen Entwicklungs- und Erwerbstheorien und den Rahmenplänen Grundschule für Mecklenburg-Vorpommern. Die Verfahren werden mit Herrn Prof. Dr. Hartke von der Universität Rostock im Rahmen meiner Promotion entwickelt und evaluiert.

Ziel der Verfahren ist es, durch mehrmalige Messwiederholungen, die Lernentwicklung und den Lernfortschritt im Rechtschreib-, Lese- und Mathematikeerwerb zu dokumentieren. Mit diesen Verfahren soll beurteilt werden, welche Lernschritte Schüler bereits bewältigt haben und welche nicht. Veränderungen im Sinne von individuellen Lernzuwächsen oder Lernstagnationen sollen sichtbar werden. Ein frühzeitige Erkennen von Entwicklungsrückständen soll „einen ersten Beitrag zur Verhinderung von Minderleistungen“ (Diehl & Hartke, 2007, S. 209) leisten. Mit den Verfahren soll nach Abschluss ihrer Entwicklung dieses präventive Handeln initiiert und Lernschwierigkeiten vorgebeugt werden.

Indem Sie die diagnostischen Verfahren in Ihrer Klasse durchführen, leisten Sie einen wichtigen Beitrag zur Entwicklung von Diagnoseverfahren zur Prävention von Lernschwierigkeiten. Ich danke Ihnen ganz herzlich für Ihre Unterstützung!

Haben Sie Fragen oder Anregungen, melden Sie sich bitte bei mir:

Kristin Hinrichsen

Neusser Wall 36, 50668 Köln

Telefon: 0221/29818640 oder 015775336426

E-Mail: kristinhinrichsen@gmx.de

Inhaltsverzeichnis

1. Erprobungszeiträume.	1
2. FE-RS 2.	2
2.1 Durchführungsdauer.	2
2.2 Testmaterial.	2
2.3 Durchführungshinweise	2
2.3.1 Allgemeine Hinweise	2
2.3.2 Wörtliche Instruktionen	3
2.3.2.1 Vor dem Austeilen der Schülerhefte	3
2.3.2.2 Nach dem Austeilen der Schülerhefte.	3
2.3.2.3 Die Durchführung der FE-RS 2	4
2.3.2.4 Nach der Durchführung	6
3. FE-L 2	8
3.1 Durchführungsdauer	8
3.2 Testmaterial	8
3.3 Durchführungshinweise	8
3.3.1 Allgemeine Hinweise	8
3.3.2 Nach der Durchführung	9
4. FE-AF 2	10
4.1 Durchführungsdauer	10
4.2 Testmaterial	10
4.3 Durchführungshinweise	10
4.3.1 Allgemeine Hinweise	10
4.3.2 Wörtliche Instruktionen	11
4.3.2.1 Vor dem Austeilen der Schülerhefte	11
4.3.2.2 Nach dem Austeilen der Schülerhefte.	11
4.3.2.3 Die Durchführung der FE-AF 2	11
4.3.2.4 Nach der Durchführung	12

1. Erprobungszeiträume

Die FE-RS 2 liegt in 2 Formen vor, in Test A und Test B. Die FE-RS 2 wird in der 2. Klassenstufe 4 mal durch Sie durchgeführt, wobei sich die 2 Testformen abwechseln.

Die FE-L 2 und FE-AF liegen in einer Form vor und werden in der 2. Klassenstufe 4 mal durch Sie durchgeführt.

Zusätzlich werden Sie in der 10. Schulwoche aus Vergleichsgründen den Deutschen Rechtschreibtest für das erste und zweite Schuljahr (DERET 1-2+) und den Deutschen Mathematiktest für erste Klassen (DEMAT 1+) durchführen. Zudem werden Sie in der 40. Schulwoche aus Vergleichsgründen erneut den DERET 1-2+ sowie den Deutschen Mathematiktest für zweite Klassen (DEMAT 2+) durchführen.

Für die Erprobungszeiträume ergibt sich folgende zeitliche Abfolge:

284	10. Schulwoche (02.11.2009 – 06.11.2009)	FE-RS 2, Test A FE-L 2 FE-AF 2 DERET 1-2 + DEMAT 1+
	20. Schulwoche (11.01.2010 – 15.01.2010)	FE-RS 2, Test B FE-L 2 FE-AF 2
	30. Schulwoche (22.03.2010 – 26.03.2010)	FE-RS 2, Test A FE-L 2 FE-AF 2
	40. Schulwoche (31.05.2010 – 04.06.2010)	FE-RS 2, Test B FE-L 2 FE-AF 2 DERET 1-2 + DEMAT 2+

Innerhalb der jeweiligen Erprobungszeiträume suchen Sie sich einen beliebigen Wochentag aus, an dem Sie die FE-RS 2, die FE-L 2, die FE-AF 2 sowie den DERET 1-2+ und den DEMAT 1+/DEMAT 2+

durchführen.

2. FE-RS 2

2.1 Durchführungsdauer

Es ist davon auszugehen, dass die Durchführung einschließlich Vorbereitung am Anfang der 2. Klasse ca. 30 Minuten und am Ende der 2. Klasse ca. 15 Minuten dauert.

2.2 Testmaterial

Die FE-RS 2 ist ein Gruppentest. Es ist ein diagnostisches Verfahren zur ökonomischen Testung mehrerer Schüler - einer gesamten Klasse.

Die FE-RS 2 wird im Klassenraum durchgeführt. Führen Sie das Verfahren nicht in der letzten Unterrichtsstunde und nicht im Nachmittagsunterricht durch. Für die Durchführung werden folgende Materialien benötigt:

- 1 Testheft pro Schüler,
- 2 Stifte pro Schüler, davon 1 Ersatzstift,
- 1 Testheft für den Testleiter zur Demonstration,
- 1 Durchführungsanleitung für den Testleiter.

2.3 Durchführungshinweise

2.3.1 Allgemeine Hinweise

Um eine objektive Testdurchführung zu gewährleisten, beachten Sie folgende Hinweise:

- Halten Sie sich an die Durchführungshinweise und die wörtlichen Instruktionen.
- Führen Sie die FE-RS 2 als Gruppentest mit der gesamten Klasse durch. Die FE-RS 2 wird ohne Pause durchgeführt.
- Üben Sie die zu diktierenden Wörter vorher nicht mit den Schülern.
- Diktieren Sie die Wörter nicht überdeutlich. Vermeiden Sie eine „Rechtschreibsprache“.

- Geben Sie den Schülern während der Durchführung keine Hinweise und Hilfestellungen.
- Verhindern Sie das Abschreiben vom Sitznachbarn.

2.3.2 Wörtliche Instruktionen

2.3.2.1 Vor dem Austeilen der Schülerhefte

Heute schreibt ihr ganz besondere Wörter. Denn ich möchte wissen, wie gut ihr schon schreiben könnt. Manche Wörter sind leicht, andere Wörter sind schon sehr schwer. Versucht aber so gut zu schreiben wie ihr könnt! Ihr schreibt in dieses Heft, das ich gleich austeilen werde. Zum Schreiben braucht ihr 2 Stifte, einen Bleistift und einen Ersatzstift, falls die Bleistiftmine abbricht. Legt jetzt 2 Stifte auf euren Tisch. Ich teile jetzt die Hefte aus. Keiner öffnet das Heft oder schreibt in das Heft.

2.3.2.2 Nach dem Austeilen der Schülerhefte

Schülerhefte austeilen und warten, dass jeder Schüler 2 Stifte hat und alle anderen Materialien in den Ranzen gelegt hat.

Bevor ich euch die Wörter diktiere, die ihr schreiben sollt, füllen wir gemeinsam das Deckblatt aus. In die erste Zeile schreibt ihr euren Nachnamen. In die zweite Zeile schreibt ihr euren Vornamen und kreuzt an, ob ihr ein Mädchen oder ein Junge seid. Unter euren Namen schreibt ihr euren Geburtstag, in die nächste Zeile die Klasse. In die letzte Zeile schreibt ihr das Datum von heute. Heute ist der ...

In dem Kästchen hier unten machen wir jetzt 2 Beispiele, damit ihr wisst, wie ihr schreiben sollt. Zuerst lese ich euch das Wort vor, welches ihr schreiben sollt. Dann schreibt ihr das Wort. Beispiel 1: Das Wort lautet **Mama**. Schreibt jetzt **Mama**. Das zweite Beispielwort darunter. Das Wort lautet **an**, wie in dem Satz: Ich warte an der Ampel. Schreibt jetzt **an**.

Öffnet jetzt euer Heft. Wie wir es geübt haben, schreibt ihr jedes Wort in eine neue Zeile, immer untereinander. Wenn ihr ein Wort falsch

geschrieben habt, streicht es durch und schreibt es richtig daneben. Achtet auch auf die Groß- und Kleinschreibung. Und: Schreibt nicht von eurem Nachbarn ab.

Hat noch jemand eine Frage?

2.3.2.3 Die Durchführung der FE-RS 2

Wir fangen jetzt an.

Test A – Diktiervorlage

- Nr. 1. Das Wort lautet **Oma**. Schreibt jetzt **Oma**.
- Nr. 2. Das Wort lautet **Ofen**. Schreibt jetzt **Ofen**.
- Nr. 3. Das Wort lautet **Nase**. Schreibt jetzt **Nase**.
- Nr. 4. Das Wort lautet **Maus**. Schreibt jetzt **Maus**.
- Nr. 5. Das Wort lautet **Meise**. Schreibt jetzt **Meise**.
- Nr. 6. Das Wort lautet **reden**, wie in dem Satz: Ich weiß, worüber sie reden. Schreibt jetzt **reden**.
- Nr. 7. Das Wort lautet **Telefon**. Schreibt jetzt **Telefon**.
- Nr. 8. Das Wort lautet **Krone**. Schreibt jetzt **Krone**.
- Nr. 9. Das Wort lautet **fragen**, wie in dem Satz: Wir **fragen** nach dem Weg. Schreibt jetzt **fragen**.
- Nr. 10. Das Wort lautet **schlafen**, wie in dem Satz: Ich gehe **schlafen**. Schreibt jetzt **schlafen**.
- Nr. 11. Das Wort lautet **flach**, wie in dem Satz: Er streckt die Hand **flach** aus. Schreibt jetzt **flach**.
- Nr. 12. Das Wort lautet **hart**, wie in dem Satz: Die Steine sind **hart**. Schreibt jetzt **hart**.
- Nr. 13. Das Wort lautet **Wolke**. Schreibt jetzt **Wolke**.
- Nr. 14. Das Wort lautet **helfen**, wie in dem Satz: Wir **helfen** uns bei den Hausaufgaben. Schreibt jetzt **helfen**.
- Nr. 15. Das Wort lautet **Stern**. Schreibt jetzt **Stern**.
- Nr. 16. Das Wort lautet **Speise**. Schreibt jetzt **Speise**.
- Nr. 17. Das Wort lautet **Qualm**. Schreibt jetzt **Qualm**.
- Nr. 18. Das Wort lautet **Leute**. Schreibt jetzt **Leute**.
- Nr. 19. Das Wort lautet **singen**, wie in dem Satz: Wir **singen** ein Lied. Schreibt jetzt **singen**.

- Nr. 20. Das Wort lautet **Bank**. Schreibt jetzt **Bank**.
 Nr. 21. Das Wort lautet **Sieb**. Schreibt jetzt **Sieb**.
 Nr. 22. Das Wort lautet **zahn**, wie in dem Satz: Meine Hund beißt nicht, er ist **zahn**. Schreibt jetzt **zahn**.
 Nr. 23. Das Wort lautet **nett**, wie in dem Satz: Mein Bruder ist **nett**. Schreibt jetzt **nett**.
 Nr. 24. Das Wort lautet **lieb**, wie in dem Satz: Sie ist **lieb**. Schreibt jetzt **lieb**.
 Nr. 25. Das Wort lautet **Katze**. Schreibt jetzt **Katze**.
 Nr. 26. Das Wort lautet **backen**, wie in dem Satz: Wir **backen** viele Plätzchen. Schreibt jetzt **backen**.
 Nr. 27. Das Wort lautet **Bäume**. Schreibt jetzt **Bäume**.
 Nr. 28. Das Wort lautet **vor**, wie in dem Satz: Ich stehe **vor** dem Haus. Schreibt jetzt **vor**.
 Bitte blättert jetzt um. Wir sind auf Seite 3.
 Nr. 29. Das Wort lautet **Hüte**. Schreibt jetzt **Hüte**.
 Nr. 30. Das Wort lautet **Vase**. Schreibt jetzt **Vase**.
 Nr. 31. Das Wort lautet **Frühblüher**. Schreibt jetzt **Frühblüher**.

Test B – Diktiervorlage

- Nr. 1. Das Wort lautet **Opa**. Schreibt jetzt **Opa**.
 Nr. 2. Das Wort lautet **Esel**. Schreibt jetzt **Esel**.
 Nr. 3. Das Wort lautet **Name**. Schreibt jetzt **Name**.
 Nr. 4. Das Wort lautet **Baum**. Schreibt jetzt **Baum**.
 Nr. 5. Das Wort lautet **Seife**. Schreibt jetzt **Seife**.
 Nr. 6. Das Wort lautet **legen**, wie in dem Satz: Wir **legen** das Buch weg. Schreibt jetzt **legen**.
 Nr. 7. Das Wort lautet **Tomaten**. Schreibt jetzt **Tomaten**.
 Nr. 8. Das Wort lautet **Blume**. Schreibt jetzt **Blume**.
 Nr. 9. Das Wort lautet **tragen**, wie in dem Satz: Wir **tragen** den Einkauf nach Hause. Schreibt jetzt **tragen**.
 Nr. 10. Das Wort lautet **schlagen**, wie in dem Satz: Wir **schlagen** uns nicht. Schreibt jetzt **schlagen**.
 Nr. 11. Das Wort lautet **frech**, wie in dem Satz: Du bist **frech** gewesen. Schreibt jetzt **frech**.
 Nr. 12. Das Wort lautet **kalt**, wie in dem Satz: Draußen ist es **kalt**.

- Schreibt jetzt **kalt**.
 Nr. 13. Das Wort lautet **Hefte**. Schreibt jetzt **Hefte**.
 Nr. 14. Das Wort lautet **halten**, wie in dem Satz: Wir **halten** uns an den Händen fest. Schreibt jetzt **halten**.
 Nr. 15. Das Wort lautet **Sturm**. Schreibt jetzt **Sturm**.
 Nr. 16. Das Wort lautet **Spule**. Schreibt jetzt **Spule**.
 Nr. 17. Das Wort lautet **Quark**. Schreibt jetzt **Quark**.
 Nr. 18. Das Wort lautet **Beule**. Schreibt jetzt **Beule**.
 Nr. 19. Das Wort lautet **fangen**, wie in dem Satz: Katzen **fangen** Mäuse. Schreibt jetzt **fangen**.
 Nr. 20. Das Wort lautet **Onkel**. Schreibt jetzt **Onkel**.
 Nr. 21. Das Wort lautet **Lied**. Schreibt jetzt **Lied**.
 Nr. 22. Das Wort lautet **kahl**, wie in dem Satz: Im Herbst sind die Bäume **kahl**. Schreibt jetzt **kahl**.
 Nr. 23. Das Wort lautet **nass**, wie in dem Satz: Meine Badesachen sind **nass**. Schreibt jetzt **nass**.
 Nr. 24. Das Wort lautet **tief**, wie in dem Satz: Der See ist **tief**. Schreibt jetzt **tief**.
 Nr. 25. Das Wort lautet **Tatze**. Schreibt jetzt **Tatze**.
 Nr. 26. Das Wort lautet **packen**, wie in dem Satz: Wir **packen** unsere Koffer. Schreibt jetzt **packen**.
 Nr. 27. Das Wort lautet **Mäuse**. Schreibt jetzt **Mäuse**.
 Nr. 28. Das Wort lautet **von**, wie in dem Satz: Das ist ein Foto **von** dir. Schreibt jetzt **von**.
 Bitte blättert jetzt um. Wir sind auf Seite 3.
 Nr. 29. Das Wort lautet **Züge**. Schreibt jetzt **Züge**.
 Nr. 30. Das Wort lautet **Vampir**. Schreibt jetzt **Vampir**.
 Nr. 31. Das Wort lautet **Fahrräder**. Schreibt jetzt **Fahrräder**.

Wenn alle Schüler mit dem Schreiben fertig sind:

Ihr seid jetzt fertig, das waren alle Wörter. Das habt ihr sehr gut gemacht!

2.3.2.4 Nach der Durchführung

Testhefte einsammeln. Kontrollieren Sie bitte, ob jeder Schüler seinen

Namen, die Klasse und Schule sowie das Datum auf das Deckblatt geschrieben hat. Kontrollieren Sie die Schreibungen der Schüler nicht auf Rechtschreibung. Schreiben Sie bitte auf jedes Deckblatt bei „Testleiter“ Ihren Namen. Lassen Sie die Zeile mit „Code-Nummer“ bitte frei.

Senden Sie bitte die Testhefte zusammen mit den Testheften der FE-AF 2 und den Auswertungsbögen des 1-Minute-Lesetests nach jeder Durchführung an die Adresse der Universität Rostock. Dafür liegen Briefumschläge bereit.

Nach der letzten Testdurchführung im Juni 2010 füllen Sie bitte den beigegefügtten Fragebogen aus und senden diesen zusammen mit den Testheften der letzten Testung und dem Lehrerheft an die Adresse der Universität Rostock.

3. FE-L 2

3.1 Durchführungsdauer

Für die Durchführung der FE-L 2 besteht eine Zeiteinschränkung von 1 Minute pro Schüler.

3.2 Testmaterial

Die FE-L 2 ist ein Einzeltest. Es ist ein diagnostisches Verfahren zur Testung eines Schülers. Es können nicht mehrere Schüler gleichzeitig getestet werden.

Die FE-L 2 wird im Klassenraum durchgeführt. Führen Sie das Verfahren nicht in der letzten Unterrichtsstunde und nicht im Nachmittagsunterricht durch. Für die Durchführung werden folgende Materialien benötigt:

- 1 Wörterliste für den Schüler,
- 1 Auswertungsbogen pro Schüler,
- 1 Stift für den Testleiter,
- 1 Stoppuhr oder Uhr mit Sekundenanzeiger für den Testleiter.

3.3 Durchführungshinweise

3.3.1 Allgemeine Hinweise

Um eine objektive Testdurchführung zu gewährleisten, beachten Sie folgende Hinweise:

- Halten Sie sich an die Durchführungshinweise.
- Führen Sie die FE-L 2 als Einzeltest mit jedem Schüler ihrer Klasse einzeln durch.
- Üben Sie die zu lesenden Wörter vorher nicht mit den Schülern.
- Geben Sie den Schülern während der Durchführung keine Hinweise und Hilfestellungen.

Für den 1-Minute-Lesetest legen Sie sich die **Wörterliste** und die benötigte Anzahl von **Wörterlisten-Auswertungsbögen** für die Schüler bereit. Schreiben Sie Name, Klasse sowie Datum und Anschrift der

Schule auf die Wörterlisten-Auswertungsbögen. Legen Sie sich Ihre Uhr mit **Sekundenanzeige** oder Ihre **Stoppuhr** bereit.

Bitten Sie den Schüler so viele Wörter wie möglich von der **Wörterliste** innerhalb 1 Minute laut vorzulesen. Der Schüler soll die Wortliste von oben nach unten lesen. Zur Kontrolle liegt vor Ihnen der Wörterliste-Auswertungsbogen. Während des lauten Vorlesens streichen Sie die falsch gelesenen Wörter durch. Selbstkorrekturen sind keine Fehler. Ist die eine Minute zu Ende, sagen Sie bitte „Stopp!“ und machen einen Strich unter das zuletzt gelesene Wort. Zählen Sie nun die Anzahl insgesamt gelesener, falsch gelesener und richtig gelesener Wörter aus. Schreiben Sie die Anzahl insgesamt gelesener, falsch gelesener und richtig gelesener Wörter auf den Wörterliste-Auswertungsbogen.

Erfahrungsgemäß gelingt die Durchführung in der Klasse, wenn die Schüler einer Stillbeschäftigung nachgehen. Nacheinander kommen die einzelnen Schüler zu dem Lehrer und lesen 1 Minute entsprechend der obigen Anweisung. Dieses Vorgehen ist vorher mit der Klasse zu besprechen, ggf. exemplarisch zu üben.

3.3.2 Nach der Durchführung

Kontrollieren Sie bitte, ob auf jedem Auswertungsbogen der Name des Schülers, die Klasse, der Name der Schule sowie das Datum notiert ist.

Senden Sie bitte die Auswertungsbögen zusammen mit den Testheften der FE-RS 2 und der FE-AF 2 nach jeder Durchführung an die Adresse der Universität Rostock. Dafür liegen Briefumschläge bereit.

4. FE-AF 2

4.1 Durchführungsdauer

Es ist davon auszugehen, dass die Durchführung einschließlich Vorbereitung 40 Minuten dauert.

4.2 Testmaterial

Die FE-AF 2 ist ein Gruppentest. Es ist ein diagnostisches Verfahren zur ökonomischen Testung mehrerer Schüler - einer gesamten Klasse.

Die FE-AF 2 wird im Klassenraum durchgeführt. Führen Sie das Verfahren nicht in der letzten Unterrichtsstunde und nicht im Nachmittagsunterricht durch. Für die Durchführung werden folgende Materialien benötigt:

- 1 Testheft pro Schüler,
- 2 Stifte pro Schüler, davon 1 Ersatzstift,
- 1 Testheft für den Testleiter zur Demonstration,
- 1 Durchführungsanleitung für den Testleiter,
- einen Klassensatz Arbeitsblätter aus dem laufenden Unterricht.

4.3 Durchführungshinweise

4.3.1 Allgemeine Hinweise

Um eine objektive Testdurchführung zu gewährleisten, beachten Sie folgende Hinweise:

- Halten Sie sich an die Durchführungshinweise und die wörtlichen Instruktionen.
- Führen Sie die FE-AF 2 als Gruppentest mit der gesamten Klasse durch. Die FE-AF 2 wird ohne Pause durchgeführt. Brechen Sie die FE-AF 2 nicht ab.
- Üben Sie die zu lösenden Aufgaben vorher nicht mit den Schülern.
- Geben Sie den Schülern während der Durchführung keine Hinweise und Hilfestellungen.
- Verhindern Sie das Abschreiben vom Sitznachbarn.

4.3.2 Wörtliche Instruktionen

4.3.2.1 Vor dem Austeilen der Schülerhefte

Heute rechnet ihr ganz besondere Aufgaben. Denn ich möchte wissen, wie gut ihr schon rechnen könnt. Manche Aufgaben sind leicht, andere Aufgaben sind schon sehr schwer. (Für 1. und 2. Testung hinzufügen: Einige Aufgaben habt ihr noch nicht im Unterricht geübt und könnt sie noch nicht lösen.) Versucht aber trotzdem so gut zu rechnen wie ihr könnt! Ihr schafft das! Ihr schreibt in dieses Heft, das ich gleich austeilen werde. Zum Schreiben braucht ihr 2 Stifte, einen Bleistift und einen Ersatzstift, falls die Bleistiftmine abbricht. Legt jetzt 2 Stifte auf euren Tisch. Ich teile jetzt die Hefte aus. Keiner öffnet das Heft oder schreibt in das Heft.

4.3.2.2 Nach dem Austeilen der Schülerhefte

Schülerhefte austeilen und warten, dass jeder Schüler 2 Stifte hat und alle anderen Materialien in den Ranzen gelegt hat.

289

Bevor ihr mit dem Rechnen beginnt, füllen wir gemeinsam das Deckblatt aus. In die erste Zeile schreibt ihr euren Vornamen und kreuzt an, ob ihr ein Mädchen oder ein Junge seid. In die zweite Zeile schreibt ihr euren Nachnamen. Unter euren Namen schreibt ihr euren Geburtstag, in die nächste Zeile die Klasse. In die letzte Zeile schreibt ihr das Datum von heute. Heute ist der ...

4.3.2.3 Die Durchführung der FE-AF 2

Heute rechnet ihr eine ganze Unterrichtsstunde alleine. Lest euch jede Aufgabe durch und überlegt, was ihr machen müsst. Strengt euch bitte an und arbeitet ohne Pausen. Die Zeit zum Rechnen ist knapp, aber die Aufgaben sind zu schaffen, wenn ihr zügig arbeitet. Wenn ihr eine Aufgabe falsch gerechnet habt, streicht sie durch und schreibt sie richtig daneben. Wenn ihr mit einer Aufgabe nicht weiterkommt, macht mit der nächsten Aufgabe weiter. Und: Schreibt nicht von eurem Nachbarn ab.

Hat noch jemand eine Frage? Fangt jetzt an.

(Nach 20 Minuten) Ihr habt jetzt noch 20 Minuten Zeit. Es wäre schön, wenn Ihr jetzt schon 7 oder 8 Aufgaben geschafft hättet. Und denkt daran, wenn ihr mit einer Aufgabe nicht weiterkommt, macht mit der nächsten Aufgabe weiter. (Für 3. und 4. Testung hinzufügen: Wer mit allen Aufgaben fertig ist, kontrolliert seine Aufgaben. Danach darf er zu mir kommen, sich ein Arbeitsblatt holen und weiterarbeiten.)

(Für 1. und 2. Testung nach 30 Minuten hinzufügen: Wenn einer von euch so gut gerechnet hat wie er kann und keine weiteren Aufgaben mehr lösen kann, kontrolliert er jetzt seine Aufgaben. Danach darf er zu mir kommen, sich ein Arbeitsblatt holen und weiterarbeiten.)

(Nach 40 Minuten) Die Unterrichtsstunde ist nun zu Ende. Beendet eure Aufgabe und schließt eure Hefte. Ihr ward sehr fleißig. Das habt ihr gut gemacht.

4.3.2.4 Nach der Durchführung

Testhefte einsammeln. Kontrollieren Sie bitte, ob jeder Schüler seinen Namen, die Klasse und Schule sowie das Datum auf das Deckblatt geschrieben hat. Kontrollieren Sie die Rechnungen der Schüler nicht. Schreiben Sie bitte auf jedes Deckblatt bei „Testleiter“ Ihren Namen. Lassen Sie die Zeile mit „Code-Nummer“ bitte frei.

Senden Sie bitte die Testhefte zusammen mit den Testheften der FE-RS 2 und den Auswertungsbögen des 1-Minute-Lesetests nach jeder Durchführung an die Adresse der Universität Rostock. Dafür liegen Briefumschläge bereit.

Formative Erfassung der Rechtschreibleistung
im 2. Schuljahr (FE-RS 2)
Universität Rostock Prof. Dr. Bodo Hartke/Kristin Hinrichsen

Schülerheft

Vorname : _____ ☐ Mädchen ☐ Junge
Nachname : _____
Geburtsdag : _____
Klasse : _____
Schule : _____
Testdatum : _____

Code-Nummer : _____
Testleiter : _____

Beispiel

1.	

2.	

RW

29.	

30.	

31.	

RW

1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	
11.	
12.	
13.	
14.	

RW

15.	
16.	
17.	
18.	
19.	
20.	
21.	
22.	
23.	
24.	
25.	
26.	
27.	
28.	

Anhang C: Lehrerfragebogen zur Einschätzung der diagnostischen Verfahren zur Lernfortschrittsmessung

Formative Erfassung der Rechtschreibleistung (FE-RS 2) und der arithmetischen Fähigkeiten (FE-AF 2)
im 2. Schuljahr, Universität Rostock Hinrichsen/Hartke

Lehrerfragebogen zur Einschätzung der diagnostischen Verfahren zur Lernfortschrittsmessung

A) Fragen zur Formativen Erfassung der Rechtschreibleistung im 2. Schuljahr (FE-RS 2)

1. Nach welcher Methode unterrichten Sie im Lernbereich Rechtschreiben?

- offener Ansatz
- lehrgangsorientierte Arbeit mit Sprachbüchern
- parallele Nutzung offener Ansätze und lehrgangsorientierter Arbeit

2. Welche Sozialformen verwenden Sie im Lernbereich Rechtschreiben?

	häufig	gelegentlich	selten	nie
lehrerzentrierter Unterricht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Offener Unterricht (z.B. Stations- und Freiarbeit, Tagesplan)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Haben Sie allgemeine Anmerkungen und Hinweise zur FE-RS 2?

.....

.....

.....

4. Haben Sie Verbesserungsvorschläge für die FE-RS 2?

.....

.....

.....

5. Wie bewerten Sie die FE-RS 2? Bitte äußern Sie sich zu folgenden Aussagen!

	Stimmt genau	Stimmt etwas	Stimmt über nicht	Stimmt nicht
Der Zeitaufwand des Verfahrens ist angemessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Art der Durchführung ist praktikabel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Wortmaterial des Verfahrens ist angemessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Nutzen des Verfahrens für Ihre pädagogische Arbeit ist groß.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie haben Interesse an einer Anwendung des Verfahrens im Unterricht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Wie bewerten die Schüler Ihrer Meinung nach die FE-RS 2? Bitte äußern Sie sich zu folgenden Aussagen!

	Stimme genau	Stimme etwas	Stimme eher nicht	Stimme nicht
Die Schüler erleben das Verfahren positiv.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Schüler weigerten oder ängstigten sich, das Verfahren durchzuführen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Schüler waren nach der Durchführung des Verfahrens erschöpft.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Schüler waren nach der Durchführung des Verfahrens enttäuscht oder demotiviert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B) Fragen zur Formativen Erfassung der arithmetischen Fähigkeiten im 2. Schuljahr (FE-AF 2)

1. Haben Sie allgemeine Anmerkungen und Hinweise zur FE-AF 2?

.....

.....

.....

2. Haben Sie Verbesserungsvorschläge für die FE-AF 2?

.....

.....

.....

3. Wie bewerten Sie die FE-AF 2? Bitte äußern Sie sich zu folgenden Aussagen!

	Stimme genau	Stimme etwas	Stimme eher nicht	Stimme nicht
Der Zeitaufwand des Verfahrens ist angemessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Art der Durchführung ist praktikabel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Aufgabenschwierigkeit des Verfahrens ist angemessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Nutzen des Verfahrens für Ihre pädagogische Arbeit ist groß.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie haben Interesse an einer Anwendung des Verfahrens im Unterricht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Wie bewerten die Schüler Ihrer Meinung nach die FE-AF 2? Bitte äußern Sie sich zu folgenden Aussagen!

	Stimme genau	Stimme etwas	Stimme eher nicht	Stimme nicht
Die Schüler erleben das Verfahren positiv.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Schüler weigerten oder ängstigten sich, das Verfahren durchzuführen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Schüler waren nach der Durchführung des Verfahrens erschöpft.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Schüler waren nach der Durchführung des Verfahrens enttäuscht oder demotiviert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Anhang E: Vordruck der E-Mail mit Bitte um Unterstützung der Untersuchung

Sehr geehrte Frau .../ Sehr geehrter Herr ...,

mein Name ist Kristin Hinrichsen. Ich habe an der Universität Rostock Sonderpädagogik studiert und promoviere seit 1,5 Jahren in diesem Fach bei Herrn Prof. Dr. Bodo Hartke. Das Thema meiner Promotion ist die Entwicklung und Evaluation eines Messverfahrens zur Erfassung früher Rechtschreibfertigkeiten in der 2. Grundschulklasse.

Wie Sie wissen, besitzt Rechtschreibkompetenz eine große Bedeutung in unserer Gesellschaft. Vielen Schülern bereitet das Rechtschreiben aber trotz häuslicher und schulischer Übungen Mühe und Schwierigkeiten. Diese Rechtschreibschwierigkeiten persistieren und wirken sich nachteilig auf die schulische und berufliche Zukunft aus. Es ist daher eine dringende Aufgabe, Kinder mit Schwierigkeiten im Rechtschreiberwerb so früh wie möglich zu identifizieren und entsprechend zu fördern. Bislang stehen Ihnen als Lehrern aber kaum Diagnoseverfahren zur Verfügung, die dem Anspruch früher und formativer Diagnostik gerecht werden. Mit der Dissertation soll diese Lücke geschlossen werden. Das Ziel der Dissertation ist die Entwicklung und Evaluation eines Diagnoseverfahrens zur Lernfortschrittsmessung im Lernbereich Rechtschreiben der 2. Klasse in der Grundschule. Ziel ist es, durch mehrmalige Messwiederholungen, die Lernentwicklung und den Lernfortschritt im Rechtschreiberwerb zu dokumentieren. Es soll beurteilt werden, welche Lernschritte Schüler im Rechtschreiberwerb bereits bewältigt haben und welche nicht. Damit kann präventives Handeln initiiert und Rechtschreibschwierigkeiten vorgebeugt werden.

Auf der Grundlage von aktuellen Theorien des Rechtschreiberwerbs und dem Rahmenplan Grundschule Deutsch für M-V habe ich dieses Messverfahren bereits entwickelt. Das Messverfahren ist als Gruppentest in Form eines Wörterdiktats mit je 31 Wörtern in aufsteigender Schwierigkeit konzipiert. Es liegt in 3 Formen (A, B, C) vor und soll in der 2. Klassenstufe 6 mal durchgeführt werden, wobei sich die 3 Testformen abwechseln (A, B, C, A, B, C). Die Testdurchführung nimmt jeweils ca. 20 Minuten in Anspruch und beeinträchtigt den Unterricht nicht.

Nun gilt es zu überprüfen, ob das Verfahren den diagnostischen Gütekriterien (Objektivität, Reliabilität, Validität) entspricht und ob das Verfahren den Lernfortschritt der Schüler dokumentieren kann. Dazu möchte ich das Messverfahren an Schulen testen und freue mich sehr, wenn auch Sie mich mit Ihrer Teilnahme unterstützen!

Für genauere Informationen und weitere Fragen schreiben Sie mir bitte eine E-Mail (kristinhinrichsen@gmx.de) oder rufen mich an unter 0221/29818640.

Ich freue mich auf eine Antwort von Ihnen!

Mit freundlichen Grüßen,
K. Hinrichsen

Anhang F: Antrag auf Genehmigung der wissenschaftlichen Untersuchung an Grundschulen



Universität Rostock
Philosophische Fakultät



Universität Rostock, Philosophische Fakultät, D-18055 Rostock, Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation
Prof. Dr. Bodo Hartke, August-Bebel-Str. 28, Telefon (0381) 498-26 79, Telefax (0381) 498-26 65, e-mail: bodo.hartke@uni-rostock.de
Kristin Hinrichsen, Neusser Wall 36, D-50668 Köln, Telefon (0221) 29818640, Mobil 015775336426, e-mail: kristinhinrichsen@gmx.de

Staatliches Schulamts Neubrandenburg
Schulamtsleiter Herr Stein
Neustrelitzer Straße 120
17033 Neubrandenburg

18.09.2009

Antrag auf Genehmigung einer wissenschaftlichen Untersuchung an Grundschulen in Mecklenburg-Vorpommern zur Vorbereitung eines Modellprojektes des BM-MV (hier im Rahmen einer Dissertation)

Sehr geehrter Herr Stein,

hiermit stelle ich einen Antrag auf Genehmigung einer wissenschaftlichen Untersuchung an Grundschulen in Mecklenburg-Vorpommern im Rahmen meiner Dissertation. Der Arbeitstitel der Dissertation lautet „Entwicklung und Evaluation eines Messverfahrens zur Erfassung früher Rechtschreibfertigkeiten im Anfangsunterricht der Grundschule“. Die Dissertation wird fachlich begleitet von Herrn Prof. Dr. Bodo Hartke an der Universität Rostock, Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation. Die in dem Vorhaben evaluierten Messverfahren sollen in den sich anschließenden Schuljahren in einem vom BM-MV getragenen Modellprojekt zur Prävention sonderpädagogischen Förderbedarfs eingesetzt werden (Arbeitstitel „Evidenzbasierte Praxis“).

Es ist eine dringende Aufgabe, Kinder mit Lernschwierigkeiten so früh wie möglich zu identifizieren. Bislang stehen Lehrern aber kaum Diagnoseverfahren zur Verfügung, die dem Anspruch früher und formativer Diagnostik gerecht werden. Mit den zu entwickelnden Verfahren, der Formativen Erfassung der Rechtschreibleistung (FE-RS 2), der Lesefertigkeit (FE-L 2) und der arithmetischen Fähigkeiten (FE-AF 2) im 2. Schuljahr, und der wissenschaftlichen Untersuchung soll diese Lücke geschlossen werden. Ziel der diagnostischen Verfahren ist, durch mehrmalige Messwiederholungen, die Lernentwicklung und den Lernfortschritt im Rechtschreib-, Lese- und Mathematikerwerb zu dokumentieren. Ein somit mögliches frühzeitiges Erkennen von Entwicklungsrückständen soll „einen ersten Beitrag zur Verhinderung von Minderleistungen“ (Diehl & Hartke, 2007, S. 209) leisten. Mit den diagnostischen Verfahren soll nach Abschluss ihrer Entwicklung dieses präventive Handeln initiiert und Lernschwierigkeiten vorgebeugt werden.

Die diagnostischen Verfahren zur Lernfortschrittsmessung sollen im Schuljahr 2009/2010 in allen 2. Grundschulklassen der Stadt Neubrandenburg und des Kreises Mecklenburg-Strelitz durch die jeweiligen Fachlehrer durchgeführt werden. Die FE-RS 2 liegt in 2 Formen vor, A und B. Die FE-RS 2 soll in der 2. Klassenstufe 4 mal im Abstand von 10 Wochen durchgeführt werden, wobei sich die 2 Testformen abwechseln. Die FE-L 2 und FE-AF liegen in parallelen Testformen vor und werden in der 2. Klassenstufe ebenfalls 4 mal durchgeführt. Zusätzlich sollen in der 10. Schulwoche der Deutsche Rechtschreibtest für das erste und zweite Schuljahr (DERET 1-2+) und der Deutsche Mathematiktest für erste Klassen (DEMAT 1+) durch die Lehrer durchgeführt werden, um die externe Validität der Messverfahren zu überprüfen. Zudem sollen in der 40. Schulwoche erneut der DERET 1-2+ sowie der Deutsche Mathematiktest für zweite Klassen (DEMAT 2+) durch die Lehrer durchgeführt werden. Daraus ergibt sich für die Erprobungszeiträume folgende zeitliche Abfolge:

10. Schulwoche 02.11.2009-06.11.2009	20. Schulwoche 11.01.2010-15.01.2010	30. Schulwoche 22.03.2010-26.03.2010	40. Schulwoche 31.05.2010-04.06.2010
FE-RS 2 FE-L 2 FE-AF 2 DERET 1-2 + DEMAT 1+	FE-RS 2 FE-L 2 FE-AF 2	FE-RS 2 FE-L 2 FE-AF 2	FE-RS 2 FE-L 2 FE-AF 2 DERET 1-2 + DEMAT 2+

Die FE-RS 2 nimmt inklusive Vor- und Nachbereitung ca. 20 Minuten in Anspruch. Die FE-L 2 dauert pro Schüler 1 Minute. Die FE-AF 2 nimmt inklusive Vor- und Nachbereitung 45 Minuten in Anspruch. Die Lehrkräfte erhalten eine Durchführungsanweisung. Die Schüler erhalten Testhefte. Im Anschluss an jede Durchführung sollen die Schülerhefte an das Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation der Universität Rostock per Post gesendet werden. Dazu liegen vorbereitete Briefumschläge vor. Im Anschluss an die Untersuchung sollen die Lehrer aufgefordert werden, einen Fragebogen zur Akzeptanz der Messverfahren zu beantworten.

Alle erhobenen Daten werden entsprechend des Datenschutzes anonymisiert und archiviert. Die personenbezogenen Daten werden nur zu wissenschaftlichen Zwecken, wie beantragt, genutzt und verarbeitet.

Nach Abschluss der Datenauswertung (Dezember 2010) werden die Ergebnisse der Untersuchung den Schulen zur Verfügung gestellt.

Mit freundlichen Grüßen,

Kristin Hinrichsen

Prof. Dr. Bodo Hartke

Anlage

Lehrerheft - Durchführungsanleitungen

Schülerheft für FE-RS 2 und FE-AF 2

Wörterliste und Wörterliste-Auswertungsbogen für FE-L 2

Fragebogen

Kopie des DERET 1-2+, des DEMAT 1+ und des DEMAT 2+

Anhang G: Mitteilung über die Antragsstellung



Universität Rostock | Philosophische Fakultät | Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation 18055 Rostock, August-Bebel-Straße 28

INSTITUT FÜR
SONDERPÄDAGOGISCHE
ENTWICKLUNGSFÖRDERUNG
UND REHABILITATION

Prof. Dr. sc. paed. habil. Bodo Hartke

Pädagogik bei
Lernbeeinträchtigungen/
Lernbehindertenpädagogik

Fon +49(0)381 498-2679
Fax +49(0)381 498-2665

bodo.hartke@uni-rostock.de

Adresse der Schule

20.09.2009

Mitteilung über Antragsstellung auf Genehmigung einer wissenschaftlichen Untersuchung an Grundschulen in Mecklenburg-Vorpommern zur Vorbereitung eines Modellprojektes des BM-MV (hier im Rahmen einer Dissertation)

Sehr geehrter Frau .../Sehr geehrter Herr ...,

hiermit teile ich Ihnen mit, dass ich beim Schulamt Neubrandenburg einen Antrag auf Genehmigung einer wissenschaftlichen Untersuchung in allen 2. Grundschulklassen der Stadt Neubrandenburg und des Kreises Mecklenburg-Strelitz im Rahmen meiner Dissertation gestellt habe. Der Arbeitstitel meiner Dissertation lautet „Entwicklung und Evaluation eines Messverfahrens zur Erfassung früher Rechtschreibfertigkeiten im Anfangsunterricht der Grundschule“. Die Dissertation wird fachlich begleitet von Herrn Prof. Dr. Bodo Hartke an der Universität Rostock, Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation. Die in dem Vorhaben evaluierten Messverfahren sollen in den sich anschließenden Schuljahren in einem vom BM-MV getragenen Modellprojekt zur Prävention sonderpädagogischen Förderbedarfs eingesetzt werden (Arbeitstitel „Evidenzbasierte Praxis“).

Es ist eine dringende Aufgabe, Kinder mit Lernschwierigkeiten so früh wie möglich zu identifizieren. Bislang stehen Ihnen als Lehrern aber kaum Diagnoseverfahren zur Verfügung, die dem Anspruch früher und formativer Diagnostik gerecht werden. Mit den zu entwickelnden Verfahren, der Formativen Erfassung der Rechtschreibleistung (FE-RS 2), der Lesefertigkeit (FE-L 2) und der arithmetischen Fähigkeiten (FE-AF 2) im 2. Schuljahr, und der wissenschaftlichen Untersuchung soll diese Lücke geschlossen werden. Ziel der diagnostischen Verfahren ist, durch mehrmalige Messwiederholungen, die Lernentwicklung und den Lernfortschritt im Rechtschreib-, Lese- und Mathematikerwerb zu dokumentieren. Ein frühzeitiges Erkennen von Entwicklungsrückständen soll „einen ersten Beitrag zur Verhinderung von Minderleistungen“ (Diehl & Hartke, 2007, S. 209) leisten. Mit den diagnostischen Verfahren

INSTITUT FÜR SONDERPÄDAGOGISCHE ENTWICKLUNGSFÖRDERUNG UND REHABILITATION

Universität Rostock | D 18051 Rostock | Fon + 49 (0)381 498-2676 | Fax + 49 (0)381 498-2665
US-IdNr.: DE 137 385 436 | Bankverbindung Inlandzahlungen: BkR Rostock, Konto 140 015 18, BLZ 130 000 00
Bankverbindung Auslandzahlungen: IBAN: DE26 1300 0000 0014 0015 18, BIC: MARKDEF1130 | www.uni-rostock.de

soll nach Abschluss ihrer Entwicklung dieses präventive Handeln initiiert und Lernschwierigkeiten vorgebeugt werden.

Die diagnostischen Verfahren zur Lernfortschrittsmessung sollen durch Sie im Schuljahr 2009/2010 4 mal in allen 2. Grundschulklassen durchgeführt werden. Zusätzlich sollen Sie den Deutschen Rechtschreibtest für das erste und zweite Schuljahr (DERET 1-2+) und den Deutschen Mathematiktest für erste Klassen (DEMAT 1+) durchführen, um die externe Validität des Messverfahrens zu überprüfen. Zudem sollen in der 40. Schulwoche erneut durch Sie der DERET 1-2+ sowie der Deutsche Mathematiktest für zweite Klassen (DEMAT 2+) durchgeführt werden. Daraus ergibt sich für die Erprobungszeiträume folgende zeitliche Abfolge:

10. Schulwoche 02.11.2009-06.11.2009	20. Schulwoche 11.01.2010-15.01.2010	30. Schulwoche 22.03.2010-26.03.2010	40. Schulwoche 31.05.2010-04.06.2010
FE-RS 2 FE-L 2 FE-AF 2 DERET 1-2 + DEMAT 1+	FE-RS 2 FE-L 2 FE-AF 2	FE-RS 2 FE-L 2 FE-AF 2	FE-RS 2 FE-L 2 FE-AF 2 DERET 1-2 + DEMAT 2+

Die FE-RS 2 nimmt inklusive Vor- und Nachbereitung ca. 20 Minuten in Anspruch. Die FE-L 2 dauert pro Schüler 1 Minute in Anspruch. Die FE-AF 2 nimmt inklusive Vor- und Nachbereitung 45 Minuten in Anspruch. Als Lehrkräfte erhalten Sie für die Durchführung eine Durchführungsanweisung. Die Schüler erhalten Testhefte. Im Anschluss an jede Durchführung sollen Sie die Schülerhefte an das Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation der Universität Rostock per Post gesendet werden. Dazu liegen vorbereitete Briefumschläge vor.

Alle erhobenen Daten werden entsprechend des Datenschutzes anonymisiert und archiviert. Die personenbezogenen Daten werden nur zu wissenschaftlichen Zwecken, wie beantragt, genutzt und verarbeitet.

Nach Abschluss der Datenauswertung (Dezember 2010) werden die Ergebnisse der Untersuchung Ihrer Schule zur Verfügung gestellt.

Ich bitte Sie, mein Forschungsvorhaben zu unterstützen. Denn indem Sie die diagnostischen Verfahren in Ihrer Schule durchführen, leisten Sie einen wichtigen Beitrag zur Entwicklung von Diagnoseverfahren zur Prävention von Lernschwierigkeiten.

Ich freue mich auf Ihre Unterstützung und die Zusammenarbeit mit Ihnen.

Mit freundlichen Grüßen,

K. Hinrichsen

Anlage

Lehrerheft - Durchführungsanleitungen

Schülerheft für FE-RS 2 und FE-AF 2

Wörterliste und Wörterliste-Auswertungsbogen für FE-L 2

**Anhang H: Bewilligung der wissenschaftlichen Untersuchung durch das
Schulamt Neubrandenburg**

**Staatliches Schulamt
Neubrandenburg**



Staatliches Schulamt Neubrandenburg
Neustrelitzer Str. 120, 17033 Neubrandenburg

Universität Rostock
Philosophische Fakultät
Herrn Prof. Dr. Hartke
A.-Bebel-Str. 28
18055 Rostock

Bearbeitet von: Junker, Susanne
Telefon: (03 95) 38 0 - 3150
e-mail: SJunker@schulamt-nb.bm.mv-regierung.de
Az: 130
Neubrandenburg, den 14. Oktober 2009

**Wissenschaftliche Untersuchung an Grundschulen der Stadt Neubrandenburg
sowie Mecklenburg-Strelitz
hier: Antrag von Frau Kristin Hinrichsen vom 18.09.2009**

Sehr geehrter Herr Prof. Dr. Hartke,

auf der Grundlage des Schulgesetzes des Landes M-V § 71 i.V.m. § 97 Abs. 1 Nr. 4 SchulG M-V und § 1 Satz 2 SchAVO M-V wird der beabsichtigten wissenschaftlichen Untersuchung zum Thema „Entwicklung und Evaluation eines Messverfahrens zur Erfassung früher Rechtschreibfertigkeiten im Anfangsunterricht der Grundschule“ an den o. g. Grundschulen stattgegeben.

Personenbezogene Daten sind nur zu wissenschaftlichen Zwecken zu nutzen und zu verarbeiten.

Die erhobenen Daten sind zu anonymisieren.

Bitte stimmen Sie die Termine rechtzeitig mit den betroffenen Schulleitern ab.

Dieses Schreiben ist bei Ihrer Kontaktaufnahme mit der ausgesuchten Schule als Genehmigungsschreiben beizufügen.

Eine gesonderte Information der Schulen durch mich erfolgt nicht.

Mit freundlichen Grüßen

i. d. g. Gust
Stein
Schulamtsleiter

Hausanschrift:
Staatliches Schulamt Neubrandenburg
Neustrelitzer Str. 120, 17033 Neubrandenburg

Telefon: +49 39 5 / 380 - 31 00
Telefax: +49 39 5 / 380 - 31 03

Universität Rostock, Philosophische Fakultät
Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation
August-Bebel-Str. 28
18055 Rostock

Ansprechpartner: Kristin Hinrichsen
E-mail: kristinhinrichsen@gmx.de

Sehr geehrte Eltern und Erziehungsberechtigte,

im Rahmen meiner Doktorarbeit möchte ich in der Klasse Ihres Kindes im Schuljahr 2009/2010 Verfahren zur Lernfortschrittsmessung im Lernbereich Lesen, Rechtschreiben und Mathematik erproben.

Das Ziel der Verfahren ist, die Lernentwicklung und die Lernschritte von Schülern im Lesen, Rechtschreiben und Rechnen zu dokumentieren. Dadurch sollen Lernschwierigkeiten frühzeitig erkannt und diesen vorgebeugt werden. Die *frühe Förderung* steht im Mittelpunkt- es wird nicht auf Störungen reagiert, sondern das Auftreten dieser verhindert. *Das Ziel ist die Optimierung des Lernens Ihres Kindes!*

Die erhobenen Daten Ihres Kindes werden entsprechend des Datenschutzes anonymisiert und archiviert. Ich bitte Sie, Ihr Einverständnis dafür zu geben, dass Ihr Kind an der Erprobung der Verfahren zur Lernfortschrittsmessung teilnehmen darf.

Vielen Dank,

Kristin Hinrichsen

✂

Einverständniserklärung

Ich bin damit *einverstanden*, dass meine Tochter/mein Sohn (Name des Kindes) an der Erprobung der Verfahren zur Lernfortschrittsmessung im Schuljahr 2009/2010 teilnimmt.

☐ Ja ☐ Nein

.....
Datum, Unterschrift Eltern/Erziehungsberechtigte

Universität Rostock, Philosophische Fakultät
Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation
August-Bebel-Str. 28
18055 Rostock

Ansprechpartner: Kristin Hinrichsen
E-mail: kristinhinrichsen@gmx.de

Sehr geehrte Eltern und Erziehungsberechtigte,

im Rahmen meiner Doktorarbeit möchte ich in der Klasse Ihres Kindes im Schuljahr 2009/2010 Verfahren zur Lernfortschrittsmessung im Lernbereich Lesen, Rechtschreiben und Mathematik erproben.

Das Ziel der Verfahren ist, die Lernentwicklung und die Lernschritte von Schülern im Lesen, Rechtschreiben und Rechnen zu dokumentieren. Dadurch sollen Lernschwierigkeiten frühzeitig erkannt und diesen vorgebeugt werden. Die *frühe Förderung* steht im Mittelpunkt- es wird nicht auf Störungen reagiert, sondern das Auftreten dieser verhindert. *Das Ziel ist die Optimierung des Lernens Ihres Kindes!*

Die erhobenen Daten Ihres Kindes werden entsprechend des Datenschutzes anonymisiert und archiviert. Ich bitte Sie, Ihr Einverständnis dafür zu geben, dass Ihr Kind an der Erprobung der Verfahren zur Lernfortschrittsmessung teilnehmen darf.

Vielen Dank,

Kristin Hinrichsen

✂

Einverständniserklärung

Ich bin damit *einverstanden*, dass meine Tochter/mein Sohn (Name des Kindes) an der Erprobung der Verfahren zur Lernfortschrittsmessung im Schuljahr 2009/2010 teilnimmt.

☐ Ja ☐ Nein

.....
Datum, Unterschrift Eltern/Erziehungsberechtigte

Anhang J: Überprüfung auf Normalverteilung des orthographischen Gesamtwerts und der Subskalen der Form A mittels Kolmogorov-Smirnov-Test

Tabelle J1: Überprüfung auf Normalverteilung des orthographischen Gesamtwerts der Form A über alle 4 Messzeitpunkte

	OG 1. MZP	OG 2. MZP	OG 3. MZP	OG 4. MZP
Kolmogorov-Smirnov-Z	1.09	1.31	1.87	1.27
P	.19	.07	.00	.08

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, OG = orthographischer Gesamtwert, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Tabelle J2: Überprüfung auf Normalverteilung der ELG-Subskala der Form A über alle 4 Messzeitpunkte

	ELG 1. MZP	ELG 2. MZP	ELG 3. MZP	ELG 4. MZP
Kolmogorov-Smirnov-Z	2.27	2.46	3.29	2.78
p	.00	.00	.00	.00

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELG = elementar graphematisches Können, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Tabelle J3: Überprüfung auf Normalverteilung der ELO-Subskala der Form A über alle 4 Messzeitpunkte

	ELO 1. MZP	ELO 2. MZP	ELO 3. MZP	ELO 4. MZP
Kolmogorov-Smirnov-Z	1.60	2.26	2.68	2.54
p	.01	.00	.00	.00

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELO = elementar orthographisches Können, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Tabelle J4: Überprüfung auf Normalverteilung der ERG-Subskala der Form A über alle 4 Messzeitpunkte

	ERG 1. MZP	ERG 2. MZP	ERG 3. MZP	ERG 4. MZP
Kolmogorov-Smirnov-Z	1.99	1.95	1.96	2.10
p	.00	.00	.00	.00

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Anhang K: Überprüfung auf Normalverteilung des orthographischen Gesamtwerts und der Subskalen der Form B mittels Kolmogorov-Smirnov-Test

Tabelle K1: Überprüfung auf Normalverteilung des orthographischen Gesamtwerts der Form B über alle 4 Messzeitpunkte

	OG 1. MZP	OG 2. MZP	OG 3. MZP	OG 4. MZP
Kolmogorov-Smirnov-Z	1.04	1.62	1.37	1.84
p	.23	.01	.05	.00

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, OG = orthographischer Gesamtwert, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Tabelle K2: Überprüfung auf Normalverteilung der ELG-Subskala der Form B über alle 4 Messzeitpunkte

	ELG 1. MZP	ELG 2. MZP	ELG 3. MZP	ELG 4. MZP
Kolmogorov-Smirnov-Z	2.38	2.81	3.07	3.07
p	.00	.00	.00	.00

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELG = elementar graphematisches Können, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Tabelle K3: Überprüfung auf Normalverteilung der ELO-Subskala der Form B über alle 4 Messzeitpunkte

	ELO 1. MZP	ELO 2. MZP	ELO 3. MZP	ELO 4. MZP
Kolmogorov-Smirnov-Z	2.17	2.22	2.27	3.08
p	.00	.00	.00	.00

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELO = elementar orthographisches Können, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Tabelle K4: Überprüfung auf Normalverteilung der ERG-Subskala der Form B über alle 4 Messzeitpunkte

	ERG 1. MZP	ERG 2. MZP	ERG 3. MZP	ERG 4. MZP
Kolmogorov-Smirnov-Z	2.48	1.77	1.96	2.20
p	.00	.00	.00	.00

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ERG = erweitert graphematisches Können, MZP = Messzeitpunkt, p = asymptotische Signifikanz (2-seitig)

Anhang L: Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form A über alle 4 Messzeitpunkte

Tabelle L1: Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert der Form A über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der OG	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Oma	⟨Oma⟩	.98	.98	.98	.99	-.02	.11	.13	.14
2	Ofen	⟨Ofen⟩	.90	.89	.90	.86	.18	.16	.14	.16
3	Nase	⟨Nase⟩	.94	.95	.97	.97	.22	.20	.13	.16
4	Maus	⟨Maus⟩	.98	1	.99	.99	.12	.15	.08	.23
5	Meise	⟨Meise⟩	.89	.94	.91	.90	.18	.13	.14	.15
6	reden	⟨reden⟩	.70	.59	.76	.73	.36	.59	.49	.57
7	Telefon	⟨Telefon⟩	.79	.87	.87	.88	.26	.18	.36	.18
8	Krone	⟨Krone⟩	.88	.84	.90	.88	.32	.30	.31	.26
9	fragen	⟨fragen⟩	.62	.59	.79	.81	.50	.34	.49	.50
10	schlafen	⟨schlafen⟩	.59	.60	.72	.76	.51	.59	.50	.47
11	flach	⟨flach⟩	.64	.62	.78	.85	.38	.58	.56	.39
12	hart	⟨hart⟩	.55	.49	.68	.66	.43	.61	.51	.47
13	Wolke	⟨Wolke⟩	.69	.71	.71	.74	.30	.38	.26	.31
14	helfen	⟨helfen⟩	.67	.57	.74	.77	.36	.60	.48	.48
15	Stern	⟨Stern⟩	.77	.81	.87	.93	.39	.45	.38	.26
16	Speise	⟨Speise⟩	.65	.62	.82	.71	.36	.36	.30	.26
17	Qualm	⟨Qualm⟩	.35	.30	.40	.36	.36	.35	.32	.24
18	Leute	⟨Leute⟩	.72	.71	.80	.75	.25	.33	.19	.42
19	singen	⟨singen⟩	.44	.49	.69	.70	.53	.59	.47	.50
20	Bank	⟨Bank⟩	.52	.62	.69	.68	.35	.41	.40	.42
21	Sieb	⟨Sieb⟩	.30	.30	.48	.46	.25	.29	.35	.31
22	zahm	⟨zahm⟩	.17	.13	.26	.29	.35	.42	.40	.45
23	nett	⟨nett⟩	.13	.26	.43	.45	.37	.49	.42	.51
24	lieb	⟨lieb⟩	.41	.51	.62	.68	.50	.60	.47	.37
25	Katze	⟨Katze⟩	.62	.64	.82	.86	.37	.52	.49	.29
26	backen	⟨backen⟩	.11	.35	.46	.41	.42	.53	.52	.56
27	Bäume	⟨Bäume⟩	.66	.65	.77	.77	.38	.45	.48	.32
28	vor	⟨vor⟩	.28	.45	.58	.66	.37	.47	.48	.45
29	Hüte	⟨Hüte⟩	.79	.79	.79	.81	.31	.19	.28	.16
30	Vase	⟨Vase⟩	.39	.50	.56	.55	.41	.41	.37	.45
31	Frühlüher	⟨Frühlüher⟩	.06	.08	.27	.38	.33	.34	.29	.49

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, OG = orthographischer Gesamtwert, MZP = Messzeitpunkt

Tabelle L2: Itemschwierigkeiten und Trennschärfen für die ELG-Subskala der Form A über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der ELG	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Oma	⟨oma⟩	.98	.98	.98	.99	.03	.18	.17	.11
2	Ofen	⟨ofen⟩	.91	.89	.90	.86	.09	.25	.17	.16
3	Nase	⟨nase⟩	.95	.96	.97	.98	.10	.23	.13	.09
4	Maus	⟨au⟩	1	1	.99	1	-	-	.10	.08
5	Meise	⟨ei⟩	.94	.96	.94	.92	.14	.05	-.02	.02
6	reden	⟨reden⟩	.93	.95	.91	.95	.14	.18	.29	.24
7	Telefon	⟨telefon⟩	.81	.89	.88	.89	.25	.35	.44	.15
8	Krone	⟨krone⟩	.90	.86	.91	.88	.30	.33	.37	.29
9	fragen	⟨fr⟩	.93	.93	.96	.98	.36	.38	.47	.30
10	schlafen	⟨schl⟩	.95	.95	.96	.97	.26	.35	.23	.24
11	flach	⟨fl⟩	.95	.93	.94	.99	.20	.45	.41	.13
		⟨ch⟩	.94	.93	.96	.97	.27	.41	.50	.23
12	hart	⟨hart⟩	.75	.73	.80	.80	.25	.22	.37	.30
13	Wolke	⟨lk⟩	.90	.84	.89	.89	.31	.42	.25	.17
14	helfen	⟨lf⟩	.96	.96	.92	.97	.20	.20	.48	.24
15	Stern	⟨st⟩	.86	.88	.96	.98	.32	.33	.35	.34
16	Speise	⟨sp⟩	.77	.78	.91	.90	.41	.50	.29	.22
17	Qualm	⟨qu⟩	.51	.42	.56	.49	.22	.26	.24	.16
18	Leute	⟨eu⟩	.81	.84	.86	.88	.20	.34	.13	.25
19	singen	⟨ng⟩	.74	.80	.90	.94	.34	.41	.35	.13
20	Bank	⟨nk⟩	.55	.65	.71	.72	.32	.25	.30	.28
21	Sieb									
22	zahm	⟨z⟩	.81	.85	.90	.88	.32	.34	.25	.22
23	nett									
24	lieb									
25	Katze									
26	backen									
27	Bäume									
28	vor									
29	Hüte									
30	Vase									
31	Frühblüher	⟨fr⟩	.89	.84	.88	.92	.42	.52	.15	.28
		⟨bl⟩	.81	.84	.90	.93	.36	.37	.28	.24

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELG = elementar graphematisches Können, MZP = Messzeitpunkt, - = Diese Items haben eine Varianz von null. Da Items mit einer Varianz von null keinen Zusammenhang mit dem Gesamtergebnis des Tests aufweisen, kann die Trennschärfe nicht berechnet werden.

Tabelle L3: Itemschwierigkeiten und Trennschärfen für die ELO-Subskala der Form A über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der ELO	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Oma	⟨O⟩	1	1	1	1	-	-	-	-
2	Ofen	⟨O⟩	1	1	1	1	-	.08	-	-
3	Nase	⟨N⟩	.99	.99	1	1	.07	.05	-	.38
4	Maus	⟨M⟩	1	1	1	1	-	.38	-	.38
5	Meise	⟨M⟩	.98	1	.99	1	.04	.06	.18	.10
6	reden									
7	Telefon	⟨T⟩	.98	.99	.99	.99	.24	.19	.28	.16
8	Krone	⟨K⟩	.97	.97	.98	.99	.29	.22	.23	.24
9	fragen	⟨en⟩	.98	.97	.98	.98	.07	.29	.23	-.01
10	schlafen	⟨en⟩	.99	1	.99	.98	.10	.01	.15	.06
11	flach									
12	hart									
13	Wolke	⟨W⟩	.93	.93	.94	.94	.27	.41	.46	.29
14	helfen	⟨en⟩	.99	.99	1	.99	-.02	-.01	-	.16
15	Stern	⟨S⟩	.94	.95	.97	.98	.40	.47	.45	.38
16	Speise	⟨S⟩	.90	.81	.90	.81	.29	.30	.27	.30
17	Qualm	⟨Q⟩	.90	.85	.94	.90	.20	.40	.33	.39
18	Leute	⟨L⟩	.87	.86	.95	.85	.28	.49	.35	.32
19	singen	⟨en⟩	.98	.94	.99	.99	.04	.32	-.01	.31
20	Bank	⟨B⟩	.94	.94	.97	.94	.23	.41	.36	.36
21	Sieb	⟨S⟩	.87	.78	.86	.82	.22	.32	.40	.31
		⟨b⟩	.53	.59	.66	.69	.30	.12	.28	.22
22	zahn									
23	nett									
24	lieb	⟨b⟩	.71	.83	.88	.88	.37	.27	.11	.17
25	Katze	⟨K⟩	.99	.98	.99	.99	.12	.41	-.04	.19
26	backen	⟨en⟩	.98	1	.98	1	0	0	0	.10
27	Bäume	⟨B⟩	.99	.98	.99	.97	.18	.51	.34	.24
		⟨äu⟩	.69	.68	.81	.81	.29	.26	.27	.30
28	vor	⟨v⟩	.37	.55	.64	.70	.29	.26	.30	.28
29	Hüte	⟨H⟩	.93	.93	.96	.97	.30	.19	.38	.09
		⟨ü⟩	.93	.89	.95	.91	.05	.25	.10	.13
30	Vase	⟨V⟩	.95	.96	.97	.97	.37	.45	.34	.29
		⟨v⟩	.43	.53	.58	.60	.42	.32	.31	.32
31	Frühblüher	⟨F⟩	.80	.81	.88	.95	.21	.35	.27	.31
		⟨er⟩	.90	.86	.93	.96	.22	.13	.13	.09

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, ELO = elementar orthographisches Können, MZP = Messzeitpunkt, - = Diese Items haben eine Varianz von null. Da Items mit einer Varianz von null keinen Zusammenhang mit dem Gesamtergebnis des Tests aufweisen, kann die Trennschärfe nicht berechnet werden.

Tabelle L4: Itemschwierigkeiten und Trennschärfen für die ERG-Subskala der Form A über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der ERG	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Oma									
2	Ofen									
3	Nase									
4	Maus									
5	Meise									
6	reden									
7	Telefon									
8	Krone									
9	fragen									
10	schlafen									
11	flach									
12	hart									
13	Wolke									
14	helfen									
15	Stern									
16	Speise									
17	Qualm									
18	Leute									
19	singen									
20	Bank									
21	Sieb	⟨ie⟩	.61	.63	.79	.72	.38	.36	.33	.33
22	zahm	⟨ah⟩	.30	.23	.34	.40	.37	.46	.36	.38
23	nett	⟨tt⟩	.18	.34	.52	.51	.34	.52	.29	.40
24	lieb	⟨ie⟩	.68	.77	.87	.87	.44	.46	.42	.32
25	Katze	⟨tz⟩	.67	.68	.85	.92	.35	.44	.36	.37
26	backen	⟨ck⟩	.17	.54	.58	.58	.39	.46	.37	.45
27	Bäume									
28	vor									
29	Hüte									
30	Vase									
31	Frühlüher	⟨üh⟩	.11	.22	.43	.58	.42	.54	.27	.47

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A,
ERG = erweitert graphematisches Können, MZP = Messzeitpunkt

Anhang M: Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert und die Subskalen der Form B über alle 4 Messzeitpunkte

Tabelle M1: Itemschwierigkeiten und Trennschärfen für den orthographischen Gesamtwert der Form B über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der OG	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Opa	⟨Opa⟩	.97	.98	.99	.99	.15	.05	.14	.06
2	Esel	⟨Esel⟩	.88	.95	.95	.94	.21	.23	.18	.37
3	Name	⟨Name⟩	.87	.88	.93	.95	.20	.31	.12	.24
4	Baum	⟨Baum⟩	.97	.95	.95	.97	.19	.19	.20	.22
5	Seife	⟨Seife⟩	.87	.92	.92	.96	.20	.35	.26	.25
6	legen	⟨leden⟩	.62	.81	.78	.86	.53	.41	.51	.39
7	Tomaten	⟨Tomaten⟩	.78	.91	.87	.90	.27	.23	.30	.26
8	Blume	⟨Blume⟩	.90	.92	.95	.97	.31	.30	.18	.36
9	tragen	⟨tragen⟩	.47	.68	.68	.81	.61	.66	.60	.58
10	schlagen	⟨schlagen⟩	.46	.67	.72	.75	.54	.52	.54	.49
11	frech	⟨frech⟩	.53	.72	.69	.80	.53	.57	.51	.54
12	kalt	⟨kalt⟩	.58	.76	.73	.80	.29	.56	.49	.39
13	Hefte	⟨Hefte⟩	.75	.85	.87	.88	.61	.34	.27	.42
14	halten	⟨halten⟩	.49	.70	.70	.77	.46	.45	.56	.49
15	Sturm	⟨Sturm⟩	.44	.59	.65	.81	.36	.42	.40	.28
16	Spule	⟨Spule⟩	.54	.64	.66	.79	.37	.27	.30	.23
17	Quark	⟨Quark⟩	.25	.44	.44	.58	.42	.39	.41	.38
18	Beule	⟨Beule⟩	.65	.72	.70	.67	.42	.24	.28	.27
19	fangen	⟨fangen⟩	.36	.70	.73	.83	.62	.54	.45	.55
20	Onkel	⟨Onkel⟩	.53	.71	.72	.80	.42	.31	.36	.32
21	Lied	⟨Lied⟩	.35	.56	.61	.77	.48	.48	.42	.38
22	kahl	⟨kahl⟩	.17	.18	.21	.34	.38	.21	.35	.32
23	nass	⟨nass⟩	.28	.47	.50	.57	.51	.52	.50	.41
24	tief	⟨tief⟩	.34	.54	.57	.67	.45	.62	.54	.60
25	Tatze	⟨Tatze⟩	.32	.45	.52	.70	.48	.47	.46	.38
26	packen	⟨packen⟩	.09	.29	.40	.44	.41	.48	.54	.49
27	Mäuse	⟨Mäuse⟩	.62	.69	.77	.86	.43	.26	.33	.25
28	von	⟨von⟩	.38	.60	.65	.78	.51	.44	.42	.43
29	Züge	⟨Züge⟩	.74	.81	.80	.88	.43	.23	.33	.30
30	Vampir	⟨Vampir⟩	.14	.25	.25	.39	.37	.35	.26	.27
31	Fahrräder	⟨Fahrräder⟩	.03	.04	.06	.15	.29	.23	.33	.29

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B,
OG = orthographischer Gesamtwert. MZP = Messzeitpunkt

Tabelle M2: Itemschwierigkeiten und Trennschärfen für die ELG-Subskala der Form B über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der ELG	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Opa	⟨opa⟩	.98	.99	.99	.99	.11	0	.16	-.02
2	Esel	⟨esel⟩	.90	.95	.95	.94	.25	.25	.30	.34
3	Name	⟨name⟩	.91	.94	.96	.97	.27	.22	.16	.07
4	Baum	⟨au⟩	.98	.99	.99	.99	.17	.22	.09	.02
5	Seife	⟨ei⟩	.97	1	.97	.98	.18	-	.15	.19
6	legen	⟨legen⟩	.92	.96	.95	.97	.41	.13	.16	.13
7	Tomaten	⟨tomaten⟩	.79	.91	.87	.91	.22	.23	.27	.19
8	Blume	⟨blume⟩	.92	.92	.95	.97	.44	.42	.22	.30
9	tragen	⟨tr⟩	.91	.96	.96	.96	.43	.36	.32	.28
10	schlagen	⟨schl⟩	.93	.93	.95	.96	.40	.32	.31	.30
11	frech	⟨fr⟩	.92	.94	.92	.95	.37	.38	.24	.46
		⟨ch⟩	.94	.97	.98	.99	.32	.28	.18	.17
12	kalt	⟨kalt⟩	.82	.90	.88	.89	.15	.31	.34	.27
13	Hefte	⟨ft⟩	.96	.98	.99	.98	.11	.34	0	.26
14	halten	⟨lt⟩	.95	.96	.97	.96	.06	.22	.22	.30
15	Sturm	⟨st⟩	.74	.79	.90	.94	.49	.37	.31	.45
16	Spule	⟨sp⟩	.68	.78	.87	.93	.46	.44	.33	.18
17	Quark	⟨qu⟩	.45	.66	.60	.78	.23	.39	.14	.27
18	Beule	⟨eu⟩	.73	.80	.78	.75	.38	.24	.24	.23
19	fangen	⟨ng⟩	.58	.84	.89	.92	.40	.47	.24	.48
20	Onkel	⟨nk⟩	.57	.74	.73	.81	.35	.23	.24	.36
21	Lied									
22	kahl									
23	nass									
24	tief									
25	Tatze									
26	packen									
27	Mäuse									
28	von									
29	Züge	⟨z⟩	.89	.96	.98	.98	.29	.14	.40	.13
30	Vampir									
31	Fahrräder									

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELG = elementar graphematisches Können. MZP = Messzeitpunkt. - = Diese Items haben eine Varianz von null. Da Items mit einer Varianz von null keinen Zusammenhang mit dem Gesamtergebnis des Tests aufweisen, kann die Trennschärfe nicht berechnet werden.

Tabelle M3: Itemschwierigkeiten und Trennschärfen für die ELO-Subskala der Form B über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der ELO	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Opa	⟨O⟩	.99	.99	1	1	.19	.10	.07	-
2	Esel	⟨E⟩	.98	1	.99	.99	.13	-	.02	.10
3	Name	⟨N⟩	.96	.94	.97	.98	.12	.21	.09	.08
4	Baum	⟨B⟩	.99	.99	1	1	.19	.10	.02	-
5	Seife	⟨S⟩	.98	.97	.99	.99	.14	.30	.26	.27
6	legen									
7	Tomaten	⟨T⟩	1	.99	1	.99	.03	.08	-	.44
8	Blume	⟨B⟩	1	.99	1	1	.09	.08	-	-
9	tragen	⟨en⟩	.97	.98	.99	1	.28	.12	.19	-
10	schlagen	⟨en⟩	.97	.97	.98	.99	.21	.15	.05	.05
11	frech									
12	kalt									
13	Hefte	⟨H⟩	.93	.94	.96	.96	.27	.30	.08	.29
14	halten	⟨en⟩	.94	.98	.96	.97	.21	.08	-.02	.08
15	Sturm	⟨S⟩	.86	.88	.87	.95	.32	.47	.34	.26
16	Spule	⟨S⟩	.86	.87	.87	.93	.39	.29	.39	.33
17	Quark	⟨Q⟩	.90	.89	.90	.98	.37	.30	.44	.38
18	Beule	⟨B⟩	.93	.93	.92	.95	.23	.27	.32	.11
19	fangen	⟨en⟩	.96	.98	.99	.99	.22	.17	.09	0
20	Onkel	⟨O⟩	.96	.99	1	.99	.35	.01	-	.36
		⟨el⟩	.91	.97	.97	.99	.34	.23	.13	.09
21	Lied	⟨L⟩	.70	.77	.80	.84	.41	.47	.45	.31
		⟨d⟩	.56	.74	.74	.89	.33	.43	.40	.46
22	kahl									
23	nass									
24	tief									
25	Tatze	⟨T⟩	.86	.86	.87	.93	.24	.33	.25	.27
26	packen	⟨en⟩	.96	.97	.99	.98	.29	.08	.08	.05
27	Mäuse	⟨M⟩	1	.99	.99	1	.11	.29	.04	-
		⟨äu⟩	.64	.71	.80	.87	.34	.39	.19	.27
28	von	⟨v⟩	.48	.62	.72	.81	.31	.37	.19	.41
29	Züge	⟨Z⟩	.93	.93	.96	.99	.32	.27	.26	.44
		⟨ü⟩	.89	.91	.88	.90	.25	.12	.09	.17
30	Vampir	⟨V⟩	.95	.94	.96	.97	.34	.25	.47	.52
		⟨v⟩	.32	.48	.50	.77	.25	.33	.29	.20
31	Fahrräder	⟨F⟩	.90	.89	.93	.93	.42	.39	.29	.39
		⟨ä⟩	.13	.21	.21	.42	.24	.30	.32	.23
		⟨er⟩	.91	.95	.91	.96	.09	.21	.08	.12

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ELO = elementar orthographisches Können. MZP = Messzeitpunkt. - = Diese Items haben eine Varianz von null. Da Items mit einer Varianz von null keinen Zusammenhang mit dem Gesamtergebnis des Tests aufweisen, kann die Trennschärfe nicht berechnet werden.

Tabelle M4: Itemschwierigkeiten und Trennschärfen für die ERG-Subskala der Form B über alle 4 Messzeitpunkte

Nr.	Wort	Lupenstelle der ERG	Itemschwierigkeit p_i				Trennschärfe r_{it}			
			1. MZP	2. MZP	3. MZP	4. MZP	1. MZP	2. MZP	3. MZP	4. MZP
1	Opa									
2	Esel									
3	Name									
4	Baum									
5	Seife									
6	legen									
7	Tomaten									
8	Blume									
9	tragen									
10	schlagen									
11	frech									
12	kalt									
13	Hefte									
14	halten									
15	Sturm									
16	Spule									
17	Quark									
18	Beule									
19	fangen									
20	Onkel									
21	Lied	⟨ie⟩	.68	.82	.88	.94	.47	.42	.34	.21
22	kahl	⟨ah⟩	.26	.21	.24	.42	.43	.21	.26	.13
23	nass	⟨ss⟩	.40	.61	.60	.74	.49	.35	.28	.29
24	tief	⟨ie⟩	.56	.68	.75	.82	.55	.48	.38	.38
25	Tatze	⟨tz⟩	.39	.52	.62	.79	.37	.47	.43	.32
26	packen	⟨ck⟩	.13	.35	.44	.53	.40	.37	.37	.39
27	Mäuse									
28	von									
29	Züge									
30	Vampir									
31	Fahrräder	⟨ah⟩	.20	.33	.40	.56	.43	.37	.44	.37

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, ERG = erweitert graphematisches Können. MZP = Messzeitpunkt

Anhang N: Measure of Sample Adequacy-Koeffizienten für die Items der Form A über alle 4 Messzeitpunkte

Tabelle N1: Measure of Sample Adequacy-Koeffizienten für die Items der Form A über alle 4 Messzeitpunkte

Wort	Measure of Sample Adequacy-Koeffizient			
	1. MZP	2. MZP	3. MZP	4. MZP
Oma	.335	.477	.526	.618
Ofen	.542	.671	.494	.638
Nase	.659	.689	.511	.645
Maus	.486	.628	.408	.667
Meise	.651	.632	.694	.569
reden	.769	.897	.853	.899
Telefon	.729	.710	.684	.559
Krone	.715	.787	.595	.687
fragen	.801	.862	.847	.875
schlafen	.842	.872	.905	.835
flach	.775	.860	.846	.802
hart	.848	.919	.877	.866
Wolke	.763	.828	.771	.746
helfen	.827	.874	.820	.856
Stern	.704	.884	.747	.691
Speise	.694	.820	.710	.720
Qualm	.870	.873	.785	.761
Leute	.726	.825	.565	.812
singen	.823	.926	.861	.886
Bank	.793	.852	.801	.838
Sieb	.673	.824	.777	.747
zahn	.799	.894	.810	.867
nett	.781	.911	.860	.900
lieb	.822	.934	.770	.823
Katze	.717	.896	.863	.782
backen	.821	.915	.869	.891
Bäume	.795	.866	.834	.720
vor	.812	.885	.895	.881
Hüte	.731	.590	.820	.681
Vase	.780	.874	.823	.859
Frühlüher	.766	.862	.815	.862

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A,
MZP = Messzeitpunkt

Anhang O: Hauptkomponentenanalysen der Form A zum 1. Messzeitpunkt

Tabelle O1: Eigenwerte und erklärte Gesamtvarianz der Form A zum 1. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	5.521	17.808	17.808
2	2.183	7.043	24.852
3	1.872	6.038	30.890
4	1.592	5.135	36.025
5	1.396	4.504	40.529
6	1.278	4.124	44.653
7	1.261	4.067	48.720
8	1.196	3.860	52.579
9	1.067	3.442	56.021
10	1.040	3.356	59.378
11	.999	3.224	62.601
12	.961	3.101	65.702
13	.902	2.911	68.614
14	.866	2.793	71.407
15	.852	2.748	74.155
16	.752	2.426	76.581
17	.737	2.377	78.958
18	.705	2.273	81.232
19	.644	2.078	83.309
20	.607	1.957	85.266
21	.554	1.786	87.052
22	.514	1.659	88.712
23	.508	1.639	90.351
24	.486	1.568	91.918
25	.439	1.416	93.334
26	.420	1.356	94.691
27	.381	1.231	95.921
28	.358	1.154	97.075
29	.323	1.043	98.118
30	.310	.998	99.117
31	.274	.883	100.000

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

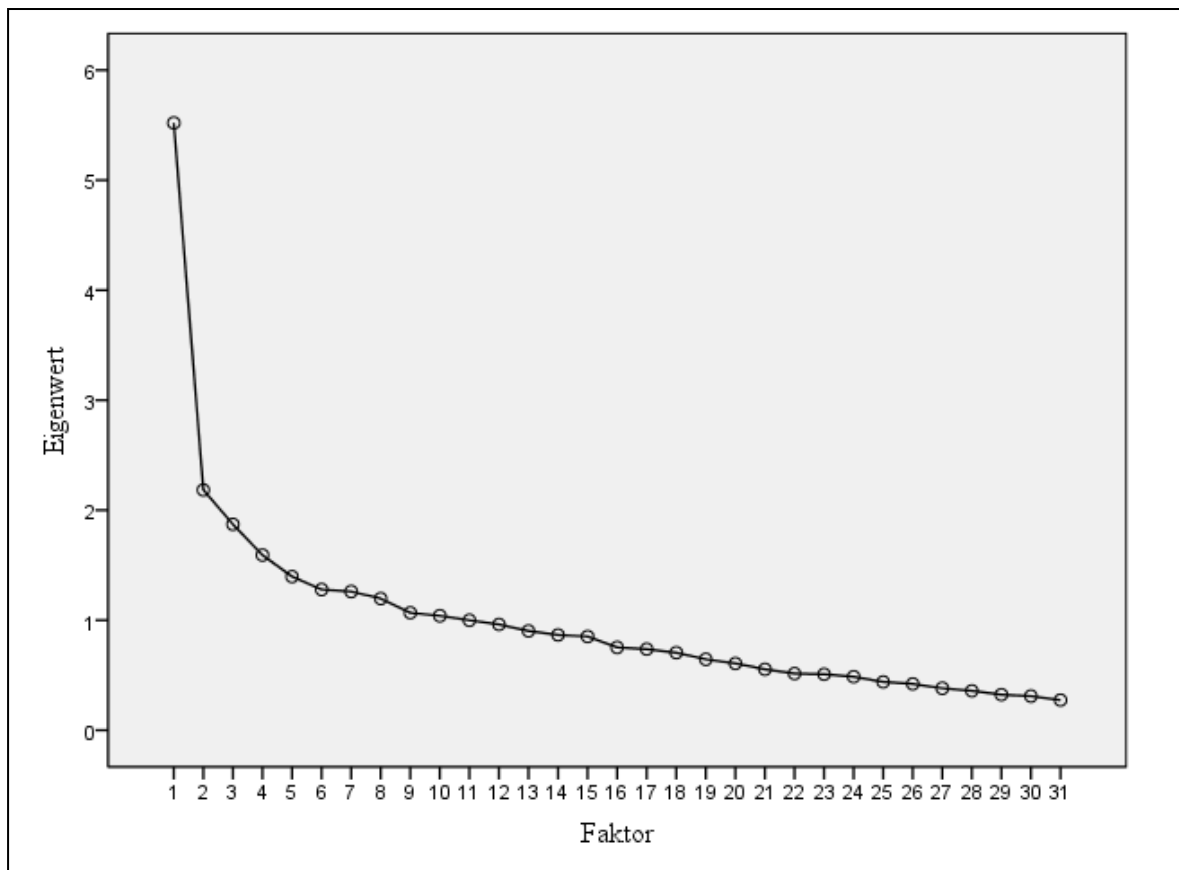


Abbildung O1: Scree-Plot der Form A zum 1. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle O2: Rotierte 10-Faktorenmatrix der Form A mit Varimax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma										.786
Ofen							.742			
Nase								.694		
Maus							.643		.375	
Meise								.698		
reden	.742									
Telefon						.732				
Krone						.696				
fragen	.750									
schlafen	.750									
flach	.638									
hart	.618									
Wolke				.408						
helfen	.357		.613							
Stern				.820						
Speise				.692						
Qualm			.302							.369
Leute									.712	
singen	.548		.371							
Bank				.327	.316					
Sieb					.797					
zahn		.308								.439
nett		.718								
lieb	.352		.378		.411					
Katze			.460			.347			-.377	
backen		.701								
Bäume			.632							
vor			.660							
Hüte							.534			
Vase					.686					
Frühlüher		.730								

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle O3: Rotierte 10-Faktorenmatrix der Form A mit Promax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma		-.385								.948
Ofen							.812			
Nase								.699		
Maus							.646		.379	
Meise								.732		
reden	.831									
Telefon						.738				
Krone						.731			.343	
fragen	.772									
schlafen	.764									
flach	.644									
hart	.602									
Wolke				.372						
helfen			.711							
Stern				.878						
Speise				.729						
Qualm										.326
Leute									.813	
singen	.478		.334							
Bank										
Sieb					.900					
zahn										.409
nett		.808								
lieb			.324		.380					
Katze			.424						-.359	
backen		.723								
Bäume			.688							
vor			.744							
Hüte							.472			
Vase					.706					
Frühlüher		.877								-.306

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle O4: Rotierte 4-Faktorenmatrix der Form A mit Varimax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma				
Ofen				.564
Nase				.545
Maus				.629
Meise				
reden	.721			
Telefon			.553	
Krone			.576	
fragen	.761			
schlafen	.745			
flach	.642			
hart	.618			
Wolke			.533	
helfen	.449			
Stern			.608	
Speise			.584	
Qualm		.546		
Leute			.406	
singen	.556	.383		
Bank			.406	
Sieb				
zahn		.508		
nett		.668		
lieb	.375	.474		
Katze		.409		
backen		.667		
Bäume		.441		
vor		.458		
Hüte			.420	.503
Vase		.376	.466	
Frühlüher		.571		

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle O5: Rotierte 4-Faktorenmatrix der Form A mit Promax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma				
Ofen				.573
Nase				.540
Maus				.642
Meise				
reden		.768		
Telefon				.576
Krone				.592
fragen		.781		
schlafen		.756		
flach		.672		
hart		.615		
Wolke				.542
helfen		.433		
Stern				.623
Speise				.601
Qualm			.563	
Leute				.406
singen		.511		
Bank				.389
Sieb				
zahn			.513	
nett			.709	
lieb			.428	
Katze			.387	
backen			.693	
Bäume			.427	
vor			.451	
Hüte				.402
Vase			.330	.446
Frühlüher			.602	

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Anhang P: Hauptkomponentenanalysen der Form A zum 2. Messzeitpunkt

Tabelle P1: Eigenwerte und erklärte Gesamtvarianz der Form A zum 2. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	7.005	22.596	22.596
2	1.989	6.416	29.012
3	1.699	5.479	34.491
4	1.541	4.970	39.461
5	1.297	4.183	43.644
6	1.251	4.034	47.678
7	1.110	3.580	51.258
8	1.070	3.451	54.710
9	1.033	3.331	58.041
10	.969	3.125	61.165
11	.928	2.993	64.159
12	.871	2.810	66.969
13	.784	2.530	69.499
14	.756	2.439	71.938
15	.736	2.375	74.313
16	.709	2.287	76.600
17	.679	2.190	78.790
18	.663	2.137	80.927
19	.605	1.953	82.880
20	.578	1.865	84.745
21	.566	1.825	86.570
22	.534	1.724	88.294
23	.500	1.613	89.907
24	.498	1.605	91.512
25	.460	1.484	92.996
26	.448	1.444	94.440
27	.433	1.398	95.838
28	.389	1.254	97.092
29	.359	1.159	98.251
30	.287	.925	99.176
31	.255	.824	100.000

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

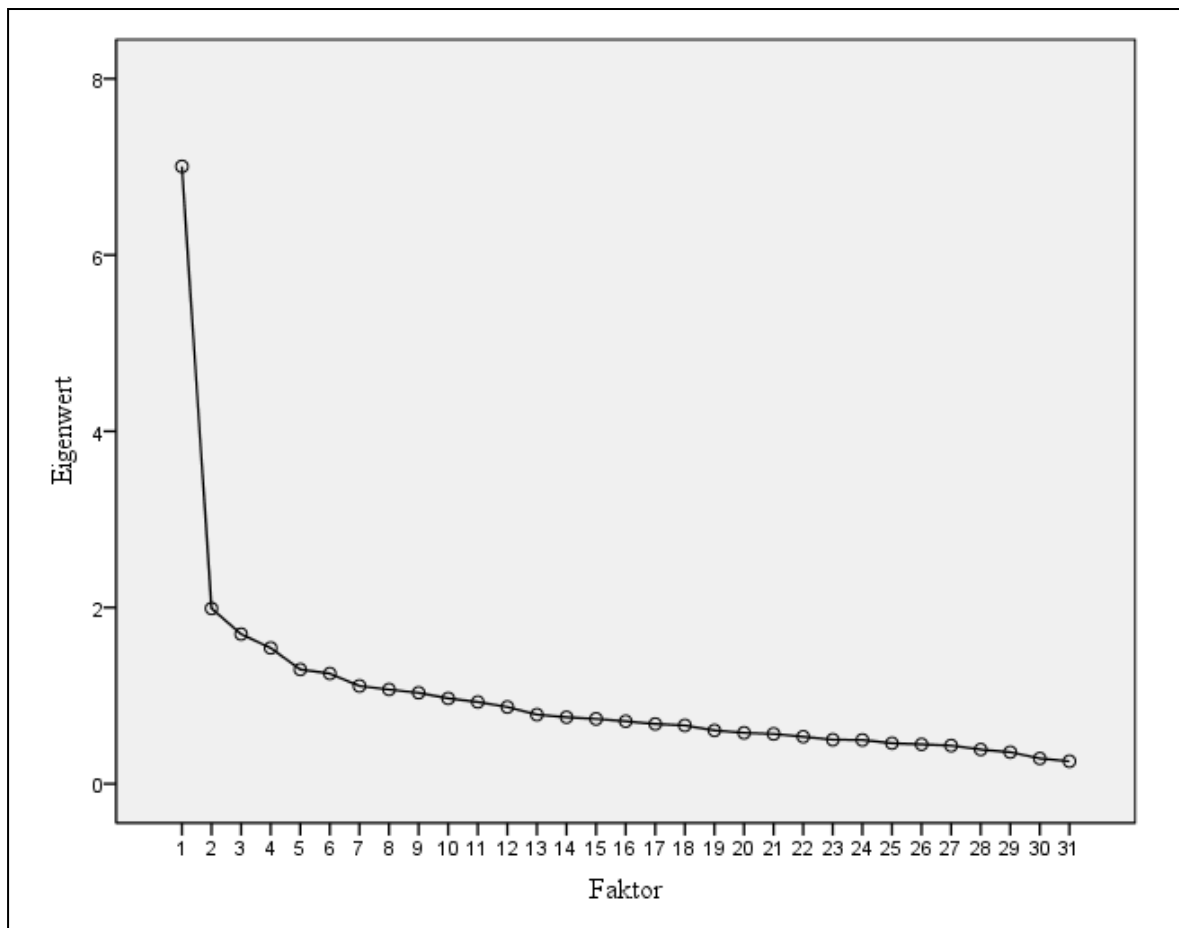


Abbildung P1: Scree-Plot der Form A zum 2. Messzeitpunkt
 Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle P2: Rotierte 9-Faktorenmatrix der Form A mit Varimax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente								
	1	2	3	4	5	6	7	8	9
Oma					.705				
Ofen							.739	.318	
Nase							.738		
Maus								.812	
Meise						.635		.445	
reden	.717								
Telefon					.466	.354			
Krone					.654				
fragen	.636								
schlafen	.788								
flach	.726								
hart	.646								
Wolke				.319	.502				
helfen	.662								
Stern				.704					
Speise				.690					
Qualm				.360					-.444
Leute						.598			
singen	.496		.463				.310		
Bank		.457		.322					
Sieb						.513			
zahn		.588							
nett		.583	.359						
Lieb	.542	.353							
Katze		.304	.456						
backen	.347	.527							
Bäume	.359		.667						
Vor		.525	.460						
Hüte									.772
Vase		.348	.497	.427					
Frühblüher		.692							

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle P3: Rotierte 9-Faktorenmatrix der Form A mit Promax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma			-.332		.780					
Ofen							.758			
Nase							.759			
Maus								.893		
Meise						.731		.537		
reden	.708									.708
Telefon					.378					
Krone					.698					
fragen	.705	-.321								.705
schlafen	.852									.852
flach	.738									.738
hart	.600									.600
Wolke					.492					
helfen	.615									.615
Stern				.775						
Speise				.771						
Qualm				.370					-.434	
Leute						.601				
singen	.386		.451							.386
Bank		.412								
Sieb						.545				
zahn		.642								
nett		.583	.325							
lieb	.445									.445
Katze			.442							
backen		.530								
Bäume			.736							
vor		.561	.467							
Hüte									.828	
Vase			.531	.391						
Frühlüher		.848								

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle P4: Rotierte 4-Faktorenmatrix der Form A mit Varimax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma			.616	
Ofen				.673
Nase				.511
Maus				.388
Meise				.461
reden	.656	.336		
Telefon			.576	
Krone			.569	
fragen	.642			
schlafen	.749			
flach	.704			
hart	.571	.461		
Wolke			.508	
helfen	.633	.386		
Stern			.379	
Speise		.313	.542	
Qualm		.424		
Leute		.351	.392	
singen	.599			.354
Bank		.487		
Sieb		.465		
zahn		.550		
nett		.626		
lieb	.543	.466		
Katze		.468		
backen		.591		
Bäume	.518			.356
vor		.561		
Hüte				.433
Vase		.478		
Frühlüher		.591		

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle P5: Rotierte 4-Faktorenmatrix der Form A mit Promax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma			.658	
Ofen				.704
Nase				.524
Maus				.393
Meise				.469
reden	.666			
Telefon			.592	
Krone			.570	
fragen	.755	-.319		
schlafen	.801			
flach	.739			
hart	.533	.342		
Wolke			.486	
helfen	.625			
Stern			.336	
Speise			.526	
Qualm		.428		
Leute		.359	.366	
singen	.597			.308
Bank		.489		
Sieb		.521		
zahn		.580		
nett		.649		
lieb	.494	.350		
Katze		.416		
backen		.582		
Bäume	.537			.333
vor		.573		
Hüte				.434
Vase		.479		
Frühlüher		.678		

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Anhang Q: Hauptkomponentenanalysen der Form A zum 3. Messzeitpunkt

Tabelle Q1: Eigenwerte und erklärte Gesamtvarianz der Form A zum 3. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	6.167	19.894	19.894
2	1.791	5.778	25.671
3	1.692	5.458	31.129
4	1.402	4.523	35.652
5	1.387	4.473	40.126
6	1.276	4.116	44.242
7	1.240	4.001	48.243
8	1.174	3.788	52.031
9	1.107	3.570	55.600
10	1.066	3.438	59.038
11	.997	3.216	62.254
12	.926	2.988	65.242
13	.868	2.799	68.041
14	.839	2.706	70.747
15	.829	2.675	73.422
16	.776	2.503	75.925
17	.737	2.379	78.304
18	.699	2.254	80.558
19	.633	2.042	82.600
20	.612	1.975	84.576
21	.594	1.918	86.493
22	.567	1.828	88.322
23	.557	1.796	90.117
24	.496	1.599	91.717
25	.441	1.423	93.140
26	.414	1.336	94.476
27	.406	1.311	95.788
28	.376	1.212	97.000
29	.352	1.135	98.135
30	.324	1.045	99.181
31	.254	.819	100.000

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

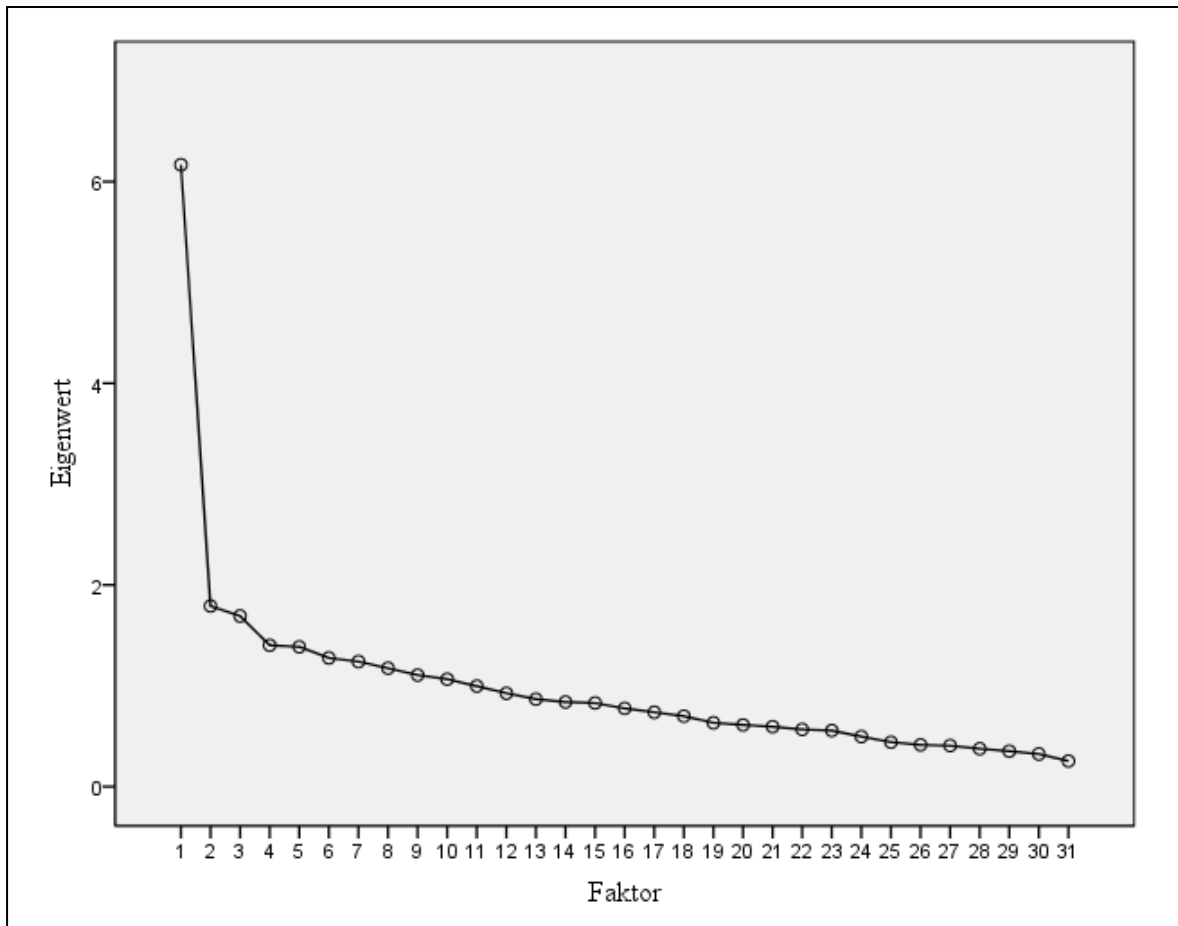


Abbildung Q1: Scree-Plot der Form A zum 3. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle Q2: Rotierte 10-Faktorenmatrix der Form A mit Varimax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma										.826
Ofen								.767		
Nase							.793			
Maus					.811					
Meise		.334				-.391		.376		
reden	.730									
Telefon				.529						
Krone				.385	.642					
fragen	.729									
schlafen	.657									
flach	.746									
hart	.501									
Wolke		.532								
helfen	.628									
Stern		.730								
Speise		.708								
Qualm									.759	
Leute				.497						-.400
singen	.538						.314			
Bank				.344			.410	-.333		
Sieb		.343	.643							
zahn						.681				
nett			.595							
lieb	.328		.643						.307	
Katze	.371	.349		.344						
backen	.346		.370			.383				
Bäume			.386	.400						
vor	.354		.326			.345				
Hüte				.628						
Vase			.381							
Frühlüher						.489			.443	

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle Q3: Rotierte 10-Faktorenmatrix der Form A mit Promax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma										.864
Ofen								.836		
Nase							.915			
Maus					.904					
Meise		.303				-.452		.398		
reden	.791									
Telefon			-.346	.583						
Krone				.323	.685	-.310				
fragen	.832									
schlafen	.733									
flach	.811									
hart	.434									
Wolke		.488								
helfen	.643									
Stern		.762					.325			
Speise		.761								
Qualm									.861	
Leute				.464					-.313	-.371
singen	.496									
Bank							.378	-.332		
Sieb			.714							
zahn						.739				
nett			.622							
lieb			.703							
Katze				.349						
backen						.343				
Bäume			.333	.391						
vor										
Hüte				.759						
Vase			.335							
Frühlüher						.528			.498	

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle Q4: Rotierte 4-Faktorenmatrix der Form A mit Varimax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma				.569
Ofen				.343
Nase				-.358
Maus			.566	
Meise		.318		.310
reden	.740			
Telefon				.481
Krone			.443	
fragen	.694			
schlafen	.622			
flach	.753			
hart	.525			
Wolke		.534		
helfen	.651			
Stern		.717		
Speise		.576		
Qualm				
Leute			.572	
singen	.564			
Bank		.519		
Sieb		.474		
zahn	.363			
nett	.344		.390	
lieb	.467	.388		
Katze	.362	.393		.343
backen	.432		.372	
Bäume			.533	
vor	.425		.357	
Hüte				.427
Vase	.318		.302	
Frühblüher			.438	

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle Q5: Rotierte 4-Faktorenmatrix der Form A mit Promax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma				.567
Ofen				.337
Nase				-.377
Maus			.645	
Meise		.357		.300
reden	.815			
Telefon				.457
Krone			.440	
fragen	.759			
schlafen	.669			
flach	.830			
hart	.504			
Wolke		.594	-.328	
helfen	.705			
Stern		.789		
Speise		.631		
Qualm				
Leute			.618	
singen	.564			
Bank		.532		
Sieb		.488		
zahn				
nett			.353	
lieb	.384	.345		-.321
Katze		.350		.306
backen	.339		.308	
Bäume			.517	
vor	.368		.308	
Hüte				.409
Vase				
Frühlüher			.439	

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Anhang R: Hauptkomponentenanalysen der Form A zum 4. Messzeitpunkt

Tabelle R1: Eigenwerte und erklärte Gesamtvarianz der Form A zum 4. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	5.970	19.259	19.259
2	1.933	6.234	25.493
3	1.580	5.097	30.590
4	1.505	4.854	35.443
5	1.353	4.364	39.807
6	1.267	4.087	43.894
7	1.178	3.801	47.695
8	1.130	3.645	51.341
9	1.061	3.423	54.764
10	1.040	3.354	58.118
11	.986	3.179	61.297
12	.935	3.017	64.315
13	.922	2.975	67.289
14	.866	2.793	70.082
15	.823	2.656	72.738
16	.770	2.485	75.223
17	.746	2.406	77.630
18	.707	2.282	79.912
19	.678	2.188	82.100
20	.609	1.964	84.064
21	.582	1.878	85.941
22	.577	1.862	87.803
23	.521	1.682	89.485
24	.482	1.554	91.039
25	.474	1.530	92.568
26	.437	1.409	93.977
27	.407	1.312	95.289
28	.402	1.297	96.586
29	.377	1.216	97.802
30	.360	1.160	98.962
31	.322	1.038	100.000

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

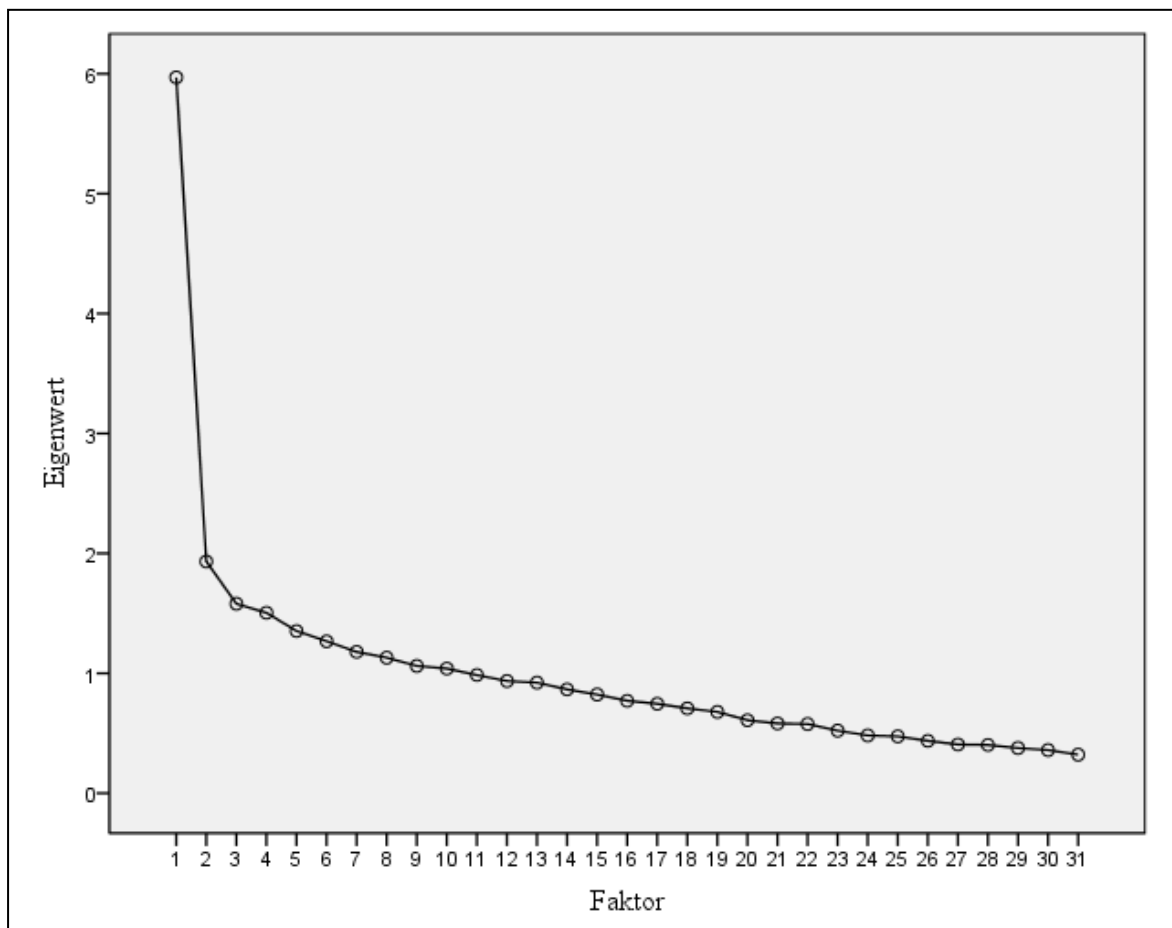


Abbildung R1: Scree-Plot der Form A zum 4. Messzeitpunkt

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle R2: Rotierte 10-Faktorenmatrix der Form A mit Varimax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma					-.311	.560		.365		
Ofen				.703						
Nase							.737			
Maus							.732			
Meise									.649	
reden	.607	.361								
Telefon				.788						
Krone			.529			.309	.413			
fragen	.766									
schlafen	.706									
flach	.729									
hart	.548									
Wolke			.791							
helfen	.538					.413				
Stern				.436		.333			.472	
Speise								.737		
Qualm		.495			-.385					
Leute			.543					.509		
singen	.545	.391								
Bank						.643				
Sieb					.599			.339		
zahn		.550								
nett		.533								
lieb		.407			.516					
Katze					.615					
backen		.694								
Bäume									.667	
vor		.535								
Hüte										.781
Vase		.523	.372							
Frühblüher		.663								

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle R3: Rotierte 10-Faktorenmatrix der Form A mit Promax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Oma					-.338	.632		.346		
Ofen				.713						
Nase							.744			
Maus							.738			
Meise									.676	
reden	.554									
Telefon				.798						
Krone			.507				.396			
fragen	.797									
schlafen	.730									
flach	.797									
hart	.521									
Wolke			.863							
helfen	.471					.426				
Stern				.395		.366			.477	
Speise								.842		.333
Qualm		.563			-.407					
Leute			.549					.476		
singen	.477	.331								
Bank						.641				
Sieb					.650			.424		
zahn		.584								
nett		.540								
lieb		.411			.461					
Katze					.645					
backen		.734								
Bäume									.685	
vor		.532								
Hüte								.321		.914
Vase		.508	.311							
Frühlüher		.723								

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle R4: Rotierte 4-Faktorenmatrix der Form A mit Varimax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma				-.329
Ofen			.527	
Nase				.588
Maus				.472
Meise				.613
reden	.631			
Telefon			.488	
Krone			.549	
fragen	.765			
schlafen	.675			
flach	.716			
hart	.568			
Wolke			.549	
helfen	.606			
Stern			.410	
Speise		.517		
Qualm		.454		
Leute		.389	.416	
singen	.595			
Bank			.416	
Sieb		.450		
zahn	.354	.387		
nett	.394	.444		
lieb	.403	.338		
Katze				.369
backen	.390	.595		
Bäume		.334		
vor	.393	.453		
Hüte				
Vase		.550		
Frühlüher		.625		

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Tabelle R5: Rotierte 4-Faktorenmatrix der Form A mit Promax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Oma				-.365
Ofen			.541	
Nase				.606
Maus				.476
Meise				.616
reden	.609			
Telefon			.494	
Krone			.536	
fragen	.841			
schlafen	.719			
flach	.831			
hart	.568			
Wolke			.527	
helfen	.638			
Stern			.405	
Speise	-.311	.599		
Qualm		.526		
Leute		.373	.378	
singen	.569			
Bank			.366	
Sieb		.492		
zahn		.356		
nett		.400		
lieb	.327	.301		
Katze				.354
backen		.590		
Bäume		.322		
vor		.428		
Hüte				
Vase		.579		
Frühlüher		.655		

Anmerkungen: Form A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A

Anhang S: Measure of Sample Adequacy-Koeffizienten für die Items der Form B über alle 4 Messzeitpunkte

Tabelle S1: Measure of Sample Adequacy-Koeffizienten für die Items der Form B über alle 4 Messzeitpunkte

Wort	Measure of Sample Adequacy-Koeffizient			
	1. MZP	2. MZP	3. MZP	4. MZP
Opa	.579	.438	.553	.452
Esel	.718	.654	.618	.796
Name	.722	.806	.493	.718
Baum	.620	.578	.627	.610
Seife	.716	.804	.643	.607
legen	.879	.792	.889	.800
Tomaten	.718	.673	.663	.689
Blume	.762	.713	.552	.700
tragen	.910	.902	.854	.888
schlagen	.898	.864	.890	.858
frech	.915	.822	.853	.832
kalt	.901	.878	.858	.774
Hefte	.667	.788	.620	.848
halten	.900	.881	.902	.850
Sturm	.853	.769	.747	.718
Spule	.823	.701	.712	.721
Quark	.897	.790	.808	.777
Beule	.831	.638	.805	.840
fangen	.916	.858	.850	.888
Onkel	.880	.838	.762	.688
Lied	.907	.816	.855	.806
kahl	.869	.679	.795	.810
nass	.870	.890	.894	.853
tief	.931	.857	.912	.902
Tatze	.845	.887	.886	.814
packen	.859	.911	.873	.887
Mäuse	.849	.705	.807	.682
von	.895	.917	.861	.779
Züge	.877	.777	.848	.706
Vampir	.850	.874	.804	.755
Fahrräder	.833	.776	.878	.805

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Anhang T: Hauptkomponentenanalysen der Form B zum 1. Messzeitpunkt

Tabelle T1: Eigenwerte und erklärte Gesamtvarianz der Form B zum 1. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	7.015	22.629	22.629
2	2.132	6.876	29.506
3	1.761	5.680	35.186
4	1.550	4.999	40.185
5	1.280	4.130	44.314
6	1.189	3.835	48.149
7	1.142	3.683	51.832
8	1.012	3.265	55.097
9	.981	3.165	58.262
10	.949	3.062	61.325
11	.905	2.918	64.243
12	.886	2.857	67.100
13	.799	2.577	69.677
14	.780	2.517	72.194
15	.735	2.370	74.565
16	.684	2.208	76.772
17	.681	2.197	78.970
18	.636	2.050	81.020
19	.628	2.027	83.047
20	.614	1.979	85.026
21	.590	1.904	86.930
22	.534	1.721	88.652
23	.505	1.630	90.281
24	.463	1.493	91.775
25	.439	1.416	93.191
26	.403	1.301	94.492
27	.380	1.226	95.717
28	.374	1.207	96.924
29	.366	1.182	98.106
30	.317	1.022	99.128
31	.270	.872	100.000

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

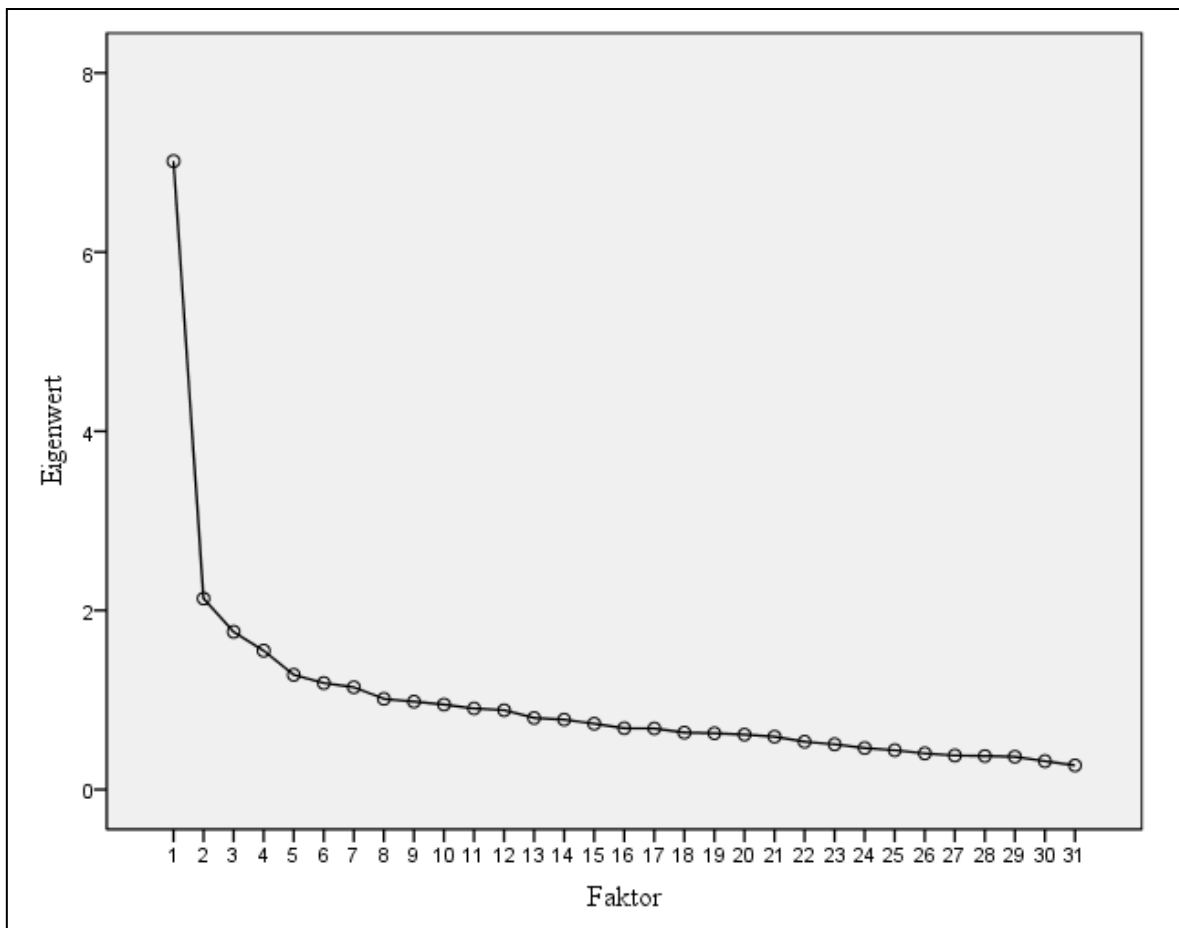


Abbildung T1: Scree-Plot der Form B zum 1. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle T2: Rotierte 8-Faktorenmatrix der Form B mit Varimax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente							
	1	2	3	4	5	6	7	8
Opa						.757		
Esel						.305		.550
Name						.375		.541
Baum						.817		
Seife								.743
legen	.774							
Tomaten							.543	
Blume				.310			.574	
tragen	.785							
schlagen	.713				.309			
frech	.656							
kalt	.564							
Hefte							.753	
halten	.727							
Sturm		.325		.388				
Spule		.346		.621				
Quark			.601					
Beule	.351			.616				
fängen	.535							
Onkel				.556				
Lied			.327		.537			
kahl		.579						
nass	.396	.634						
tief	.452	.366						
Tatze		.528						
packen			.576					
Mäuse					.688			
von			.420					
Züge				.335	.560			
Vampir		.658						
Fahrräder			.635					

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle T3: Rotierte 8-Faktorenmatrix der Form B mit Promax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente							
	1	2	3	4	5	6	7	8
Opa						.766		
Esel								.549
Name						.334		.506
Baum						.820		
Seife								.759
legen	.836							
Tomaten							.496	
Blume							.575	
tragen	.807							
schlagen	.726							
frech	.663							
kalt	.540							
Hefte							.794	
halten	.710							
Sturm				.329				
Spule				.613				
Quark			.620					
Beule	.337			.628				
fangen	.451							
Onkel				.559				
Lied					.546			
kahl		.589						
nass		.639						
tief	.401	.321						
Tatze		.473						
packen			.605					
Mäuse					.758			
von			.363					
Züge					.591			
Vampir		.689						
Fahrräder			.695					

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle T4: Rotierte 4-Faktorenmatrix der Form B mit Varimax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.660
Esel				.449
Name				.601
Baum				.757
Seife			.512	
legen	.751			
Tomaten			.365	
Blume			.644	
tragen	.761			
schlagen	.708			
frech	.655			
kalt	.583			
Hefte			.446	
halten	.720			
Sturm		.481	.376	
Spule		.491	.451	
Quark		.466		
Beule			.477	
fängen	.521	.303	.333	
Onkel			.450	
Lied			.356	
kahl		.564		
nass	.403	.600		
tief	.451	.394		
Tatze		.599		
packen		.440		
Mäuse			.339	.337
von		.536		
Züge		.349	.463	
Vampir		.574		
Fahrräder		.383		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle T5: Rotierte 4-Faktorenmatrix der Form B mit Promax-Rotation zum 1. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.689
Esel				.431
Name				.597
Baum				.776
Seife			.546	
legen	.804			
Tomaten			.368	
Blume			.695	
tragen	.772			
schlagen	.728			
frech	.669			
kalt	.577			
Hefte			.466	
halten	.718			
Sturm		.499	.326	
Spule	-.306	.511	.415	
Quark		.457		
Beule			.465	
fangen	.465			
Onkel			.423	
Lied				
kahl		.613		
nass		.614		
tief	.395	.348		
Tatze		.618		
packen		.430		
Mäuse				
von		.524		
Züge			.422	
Vampir		.629		
Fahrräder		.401		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Anhang U: Hauptkomponentenanalysen der Form B zum 2. Messzeitpunkt

Tabelle U1: Eigenwerte und erklärte Gesamtvarianz der Form B zum 2. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	6.520	21.031	21.031
2	1.927	6.216	27.247
3	1.532	4.941	32.188
4	1.391	4.486	36.674
5	1.340	4.324	40.998
6	1.241	4.005	45.003
7	1.151	3.713	48.716
8	1.109	3.578	52.294
9	1.031	3.325	55.620
10	1.017	3.280	58.900
11	.953	3.076	61.975
12	.934	3.013	64.988
13	.915	2.953	67.941
14	.842	2.716	70.657
15	.807	2.603	73.260
16	.763	2.462	75.722
17	.745	2.403	78.125
18	.714	2.302	80.427
19	.656	2.116	82.542
20	.635	2.048	84.590
21	.604	1.948	86.538
22	.556	1.793	88.331
23	.518	1.672	90.003
24	.512	1.652	91.655
25	.473	1.526	93.181
26	.449	1.448	94.629
27	.418	1.348	95.977
28	.402	1.296	97.273
29	.322	1.040	98.313
30	.282	.911	99.224
31	.241	.776	100.000

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

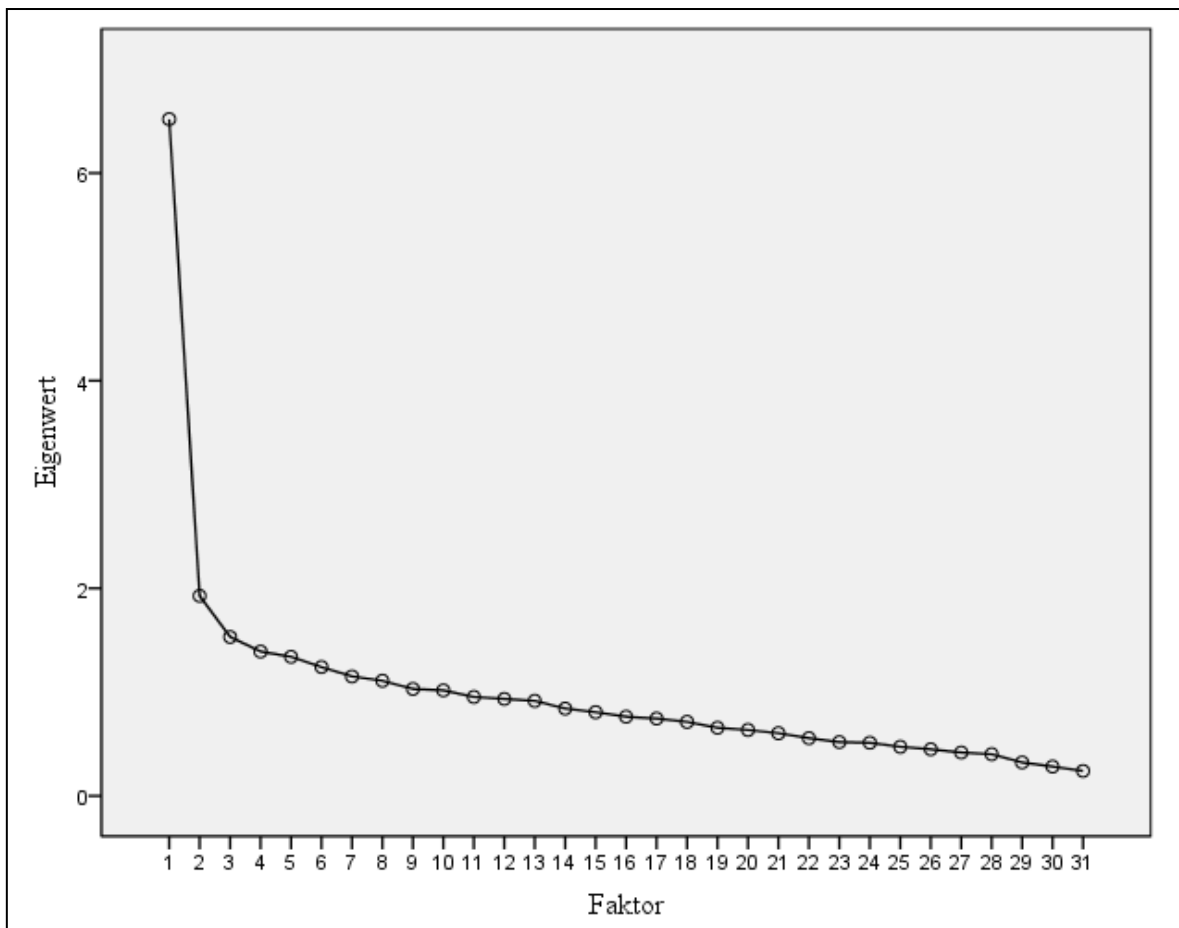


Abbildung U1: Scree-Plot der Form B zum 2. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle U2: Rotierte 10-Faktorenmatrix der Form B mit Varimax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Opa						.765				
Esel					.762					
Name					.592					
Baum								.683		
Seife					.343	.630				
legen	.649									
Tomaten			.359	.561						
Blume				.402		.607				
tragen	.727									
schlagen	.705									
frech	.730									
kalt	.670	.305								
Hefte					.605					
halten	.653									
Sturm				.651						
Spule				.761						
Quark			.658							
Beule								.512		
fängen	.588									
Onkel			.400							.420
Lied		.644								
kahl							.794			
nass	.362		.634							
tief	.395						.440			
Tatze			.533				.399			
packen	.323	.554								
Mäuse		.622								
von	.325	.493								
Züge									.822	
Vampir			.388						.306	
Fahrräder										.807

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle U3: Rotierte 10-Faktorenmatrix der Form B mit Promax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente									
	1	2	3	4	5	6	7	8	9	10
Opa						.883				
Esel					.786					
Name					.586			-.311		
Baum								.757		
Seife						.673				
legen	.729									
Tomaten			.313	.581						
Blume				.304		.594				
tragen	.729									
schlagen	.762									
frech	.758									
kalt	.677									
Hefte					.595					
halten	.725									
Sturm				.698						
Spule				.891						
Quark			.723							
Beule								.481		
fangen	.560									
Onkel			.408							.403
Lied		.649								
kahl							.840			
nass			.670							
tief							.402			
Tatze			.580				.387			
packen		.530								
Mäuse		.698								
von		.472								
Züge									.893	
Vampir			.330							-.306
Fahrräder										.834

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle U4: Rotierte 4-Faktorenmatrix der Form B mit Varimax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.612
Esel			.707	
Name			.559	
Baum				
Seife			.503	.402
legen	.617			
Tomaten		.503		
Blume		.427		.530
tragen	.715			
schlagen	.696			
frech	.737			
kalt	.655			
Hefte			.643	
halten	.648			
Sturm		.686		
Spule		.500		
Quark		.453		
Beule		.400		
fängen	.643			
Onkel		.312		
Lied		.480	.376	
kahl	.353			
nass	.460	.430		
tief	.519	.445		
Tatze	.401	.349		
packen	.330	.419		
Mäuse			.434	
von	.323	.333		
Züge				
Vampir		.323		
Fahrräder				

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle U5: Rotierte 4-Faktorenmatrix der Form B mit Promax-Rotation zum 2. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.620
Esel			.790	
Name			.575	
Baum				
Seife			.486	.415
legen	.693			
Tomaten		.603		
Blume		.461		.553
tragen	.716			
schlagen	.758			
frech	.790			
kalt	.677			
Hefte			.675	
halten	.735			
Sturm		.786		
Spule		.582		
Quark		.463		
Beule		.453		
fangen	.680			
Onkel				
Lied		.453	.306	
kahl	.424			
nass	.383	.383		
tief	.454	.362		
Tatze	.319			
packen		.376		
Mäuse			.445	
von				
Züge				
Vampir				
Fahrräder		.315		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Anhang V: Hauptkomponentenanalysen der Form B zum 3. Messzeitpunkt

Tabelle V1: Eigenwerte und erklärte Gesamtvarianz der Form B zum 3. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	6.399	20.643	20.643
2	2.214	7.141	27.784
3	1.457	4.700	32.484
4	1.407	4.537	37.022
5	1.357	4.377	41.399
6	1.198	3.864	45.262
7	1.145	3.694	48.956
8	1.129	3.641	52.597
9	1.036	3.340	55.938
10	.981	3.164	59.102
11	.949	3.062	62.163
12	.934	3.013	65.176
13	.863	2.785	67.961
14	.834	2.690	70.651
15	.819	2.643	73.294
16	.788	2.543	75.837
17	.774	2.496	78.332
18	.717	2.312	80.644
19	.654	2.108	82.753
20	.626	2.018	84.771
21	.601	1.938	86.709
22	.567	1.830	88.538
23	.522	1.685	90.223
24	.504	1.624	91.847
25	.479	1.547	93.394
26	.421	1.360	94.754
27	.402	1.296	96.050
28	.352	1.136	97.186
29	.327	1.056	98.242
30	.303	.977	99.219
31	.242	.781	100.000

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

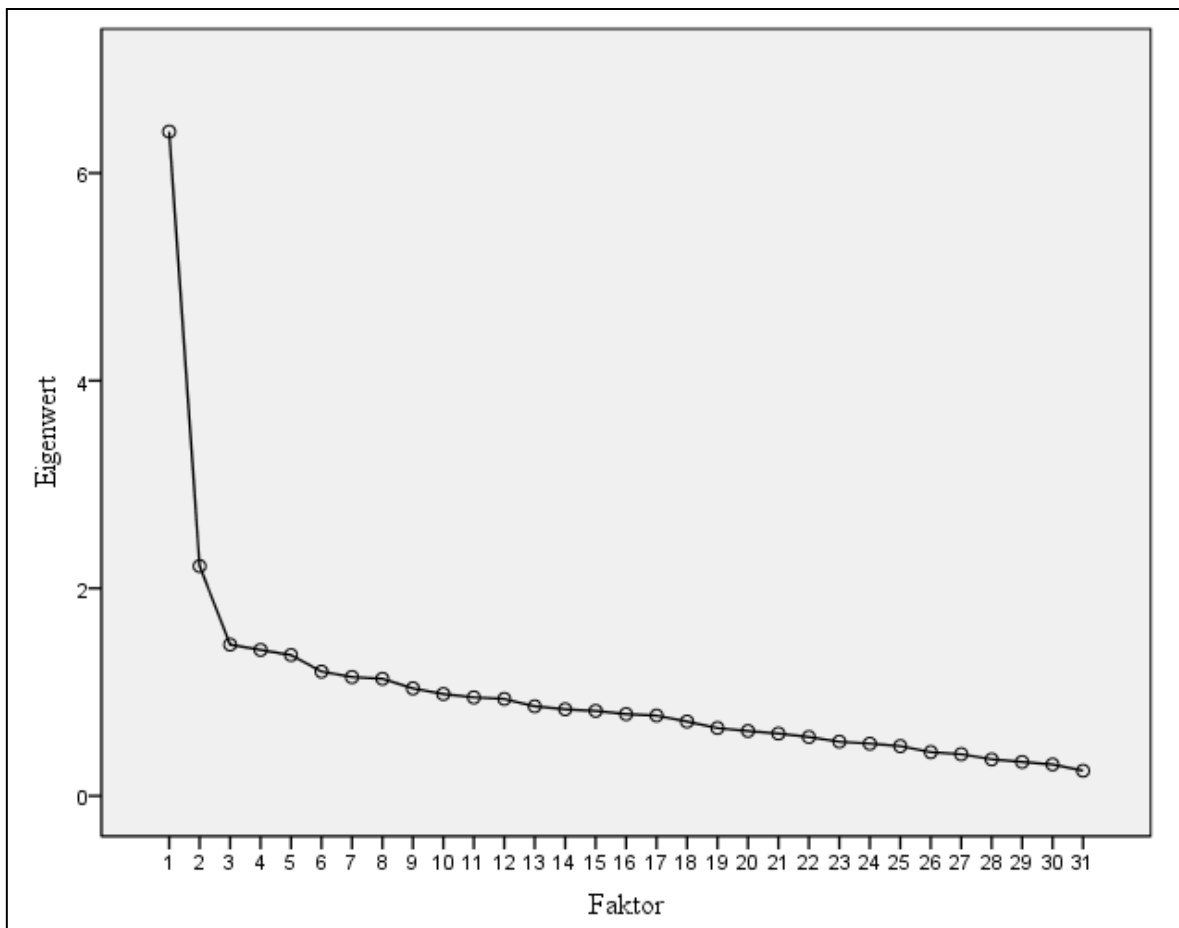


Abbildung V1: Scree-Plot der Form B zum 3. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle V2: Rotierte 9-Faktorenmatrix der Form B mit Varimax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente								
	1	2	3	4	5	6	7	8	9
Opa					.631				
Esel					.701				
Name								.810	
Baum						.685			
Seife						.581			
legen	.742								
Tomaten				.559	.400				
Blume									.782
tragen	.788								
schlagen	.791								
frech	.626								.435
kalt	.605		.350						
Hefte			-.429		.317		.498		
halten	.754								
Sturm		.739							
Spule		.543		.505					
Quark			.735						
Beule				.672					
fängen	.584								
Onkel							.506		
Lied		.451						.395	
kahl			.591						.302
nass	.362			.314					
tief	.571								
Tatze		.517					.332		
packen	.412			.340				.320	
Mäuse		.420			.471				
von	.434					.359			
Züge		.401				.332			
Vampir							.652		
Fahrräder			.350						

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle V3: Rotierte 9-Faktorenmatrix der Form B mit Promax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente								
	1	2	3	4	5	6	7	8	9
Opa					.658				
Esel					.718				
Name								.900	
Baum						.808		-.337	
Seife						.625			.314
legen	.818								
Tomaten				.605	.355				
Blume									.913
tragen	.859								
schlagen	.861								
frech	.632								.392
kalt	.577		.316						
Hefte			-.474				.463		
halten	.813								
Sturm		.835							
Spule		.622		.584					
Quark			.746						
Beule				.788					
fangen	.573								
Onkel		.300					.467		
Lied		.420						.367	
kahl			.638						.408
nass				.302					
tief	.559								
Tatze		.488					.314		
packen				.329					
Mäuse		.368			.462				
von	.384								
Züge		.353				.345			
Vampir							.740		
Fahrräder			.320						

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle V4: Rotierte 4-Faktorenmatrix der Form B mit Varimax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.601
Esel				.558
Name		.402		
Baum			.377	
Seife			.526	
legen	.734			
Tomaten				
Blume			.512	
tragen	.775			
schlagen	.790			
frech	.637		.324	
kalt	.631			
Hefte			.428	.461
halten	.738			
Sturm		.302	.385	.407
Spule			.635	
Quark		.684		
Beule			.468	
fängen	.589			
Onkel				.530
Lied		.559		
kahl		.462		
nass	.353	.398		
tief	.537			
Tatze				.374
packen	.379	.473		
Mäuse				.435
von	.430	.344		
Züge			.407	
Vampir				
Fahrräder		.472		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle V5: Rotierte 4-Faktorenmatrix der Form B mit Promax-Rotation zum 3. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa			-.339	.688
Esel				.609
Name		.467		
Baum			.419	
Seife			.593	
legen	.793			
Tomaten				
Blume			.583	
tragen	.808			
schlagen	.845			
frech	.663			
kalt	.641			
Hefte		-.372	.419	.421
halten	.766			
Sturm			.322	.341
Spule			.677	
Quark		.735		
Beule			.492	
fangen	.609			
Onkel				.523
Lied		.563		
kahl		.472		
nass		.345		
tief	.506			
Tatze				.308
packen		.424		
Mäuse				.425
von	.379			
Züge			.412	
Vampir				
Fahrräder		.488		

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Anhang W: Hauptkomponentenanalysen der Form B zum 4. Messzeitpunkt

Tabelle W1: Eigenwerte und erklärte Gesamtvarianz der Form B zum 4. Messzeitpunkt

Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	6.097	19.669	19.669
2	2.088	6.736	26.405
3	1.635	5.275	31.680
4	1.482	4.782	36.462
5	1.394	4.498	40.960
6	1.328	4.284	45.244
7	1.287	4.151	49.395
8	1.183	3.815	53.210
9	1.098	3.541	56.750
10	1.047	3.377	60.127
11	1.005	3.242	63.369
12	.961	3.101	66.471
13	.876	2.827	69.298
14	.830	2.678	71.975
15	.776	2.504	74.479
16	.737	2.377	76.856
17	.680	2.193	79.049
18	.664	2.141	81.190
19	.608	1.960	83.150
20	.600	1.936	85.086
21	.567	1.829	86.915
22	.528	1.704	88.619
23	.498	1.608	90.227
24	.479	1.547	91.773
25	.455	1.469	93.243
26	.447	1.443	94.685
27	.387	1.247	95.933
28	.370	1.193	97.125
29	.329	1.063	98.188
30	.291	.939	99.127
31	.271	.873	100.000

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

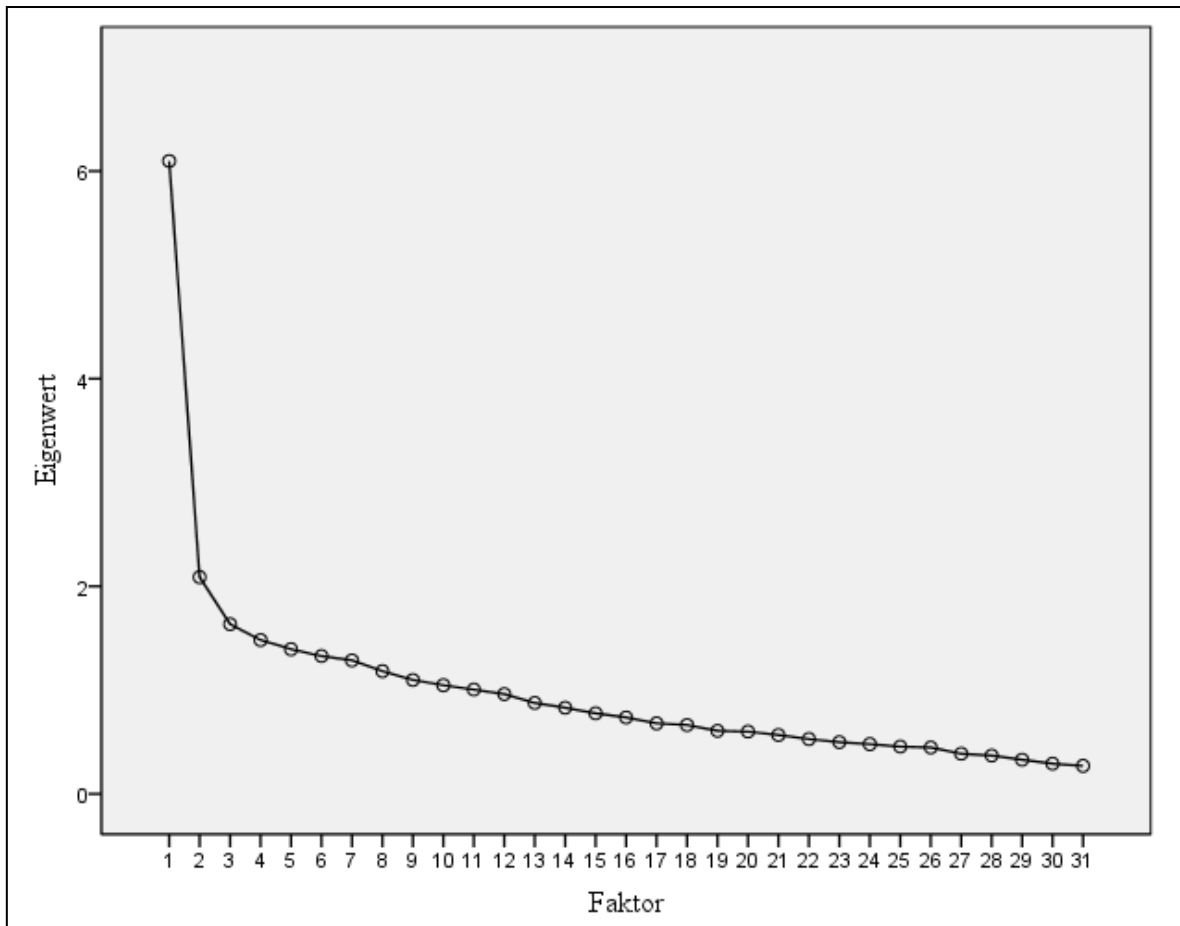


Abbildung W1: Scree-Plot der Form B zum 4. Messzeitpunkt

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle W2: Rotierte 11-Faktorenmatrix der Form B mit Varimax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente										
	1	2	3	4	5	6	7	8	9	10	11
Opa									.830		
Esel				.748							
Name					.735						
Baum						.754					
Seife				.520				.393			
legen	.632										
Tomaten						.620					
Blume					.614						
tragen	.659										
schlagen	.790										
frech	.698										
kalt	.486		.377						.332		
Hefte		.448			.416						
halten	.700										
Sturm							.397	.477			
Spule								.621			
Quark			.726								
Beule											.760
fängen	.620										
Onkel			.600	.395						.364	
Lied		.522									.567
kahl								.629			
nass	.518						.427			.315	
tief	.407	.573									
Tatze		.675									
packen	.397	.366					.381				
Mäuse		.335	.332						.453		
von		.677									
Züge				.390		.325	.350		.395		
Vampir										.761	
Fahrräder							.734				

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle W3: Rotierte 11-Faktorenmatrix der Form B mit Promax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente										
	1	2	3	4	5	6	7	8	9	10	11
Opa									.910		
Esel				.795							
Name					.801						
Baum						.827					
Seife				.568				.393			
legen	.674										
Tomaten			.358		-.315	.672					
Blume					.611						
tragen	.658										
schlagen	.898										
frech	.681										
kalt	.447		.302						.302		
Hefte		.388			.392						
halten	.772										
Sturm			.351				.404	.448			
Spule								.667			
Quark			.881								
Beule											.875
fangen	.575										
Onkel			.634	.313						.361	
Lied		.477									.535
kahl								.669			
nass	.456						.425			.339	
tief		.570									
Tatze		.741									
packen		.309					.409				
Mäuse									.422		
von		.737									
Züge				.317					.367		
Vampir										.918	
Fahrräder						.814					

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle W4: Rotierte 4-Faktorenmatrix der Form B mit Varimax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.446
Esel			.502	
Name		.438		
Baum			.465	
Seife			.522	
legen	.643		.314	
Tomaten			.489	
Blume		.511		
tragen	.614		.326	
schlagen	.738			
frech	.692	.310		
kalt	.428	.373		
Hefte		.640		
halten	.645			
Sturm			.434	
Spule		.525		
Quark			.444	
Beule				
fängen	.631	.306		
Onkel		.448	.366	
Lied		.320		.357
kahl		.325		
nass	.619			.302
tief	.427	.365		.457
Tatze		.463		.392
packen	.440			.498
Mäuse				.515
von		.424		.467
Züge			.525	.300
Vampir		.349		
Fahrräder				.477

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Tabelle W5: Rotierte 4-Faktorenmatrix der Form B mit Promax-Rotation zum 4. Messzeitpunkt (> .30)

	Komponente			
	1	2	3	4
Opa				.486
Esel			.468	
Name		.482		
Baum			.441	
Seife			.523	
legen	.701			
Tomaten			.486	
Blume		.542		
tragen	.603			
schlagen	.785			
frech	.705			
kalt	.407	.322		
Hefte		.660		
halten	.664			
Sturm			.421	
Spule		.577		
Quark			.409	
Beule				
fangen	.623			
Onkel		.483	.331	
Lied				.329
kahl				
nass	.659			
tief	.347			.394
Tatze		.428		.342
packen	.398			.468
Mäuse				.542
von		.371		.433
Züge			.522	.330
Vampir		.326		
Fahrräder				.504

Anmerkungen: Form B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Anhang X: Regressionsanalysen der FE-RS 2

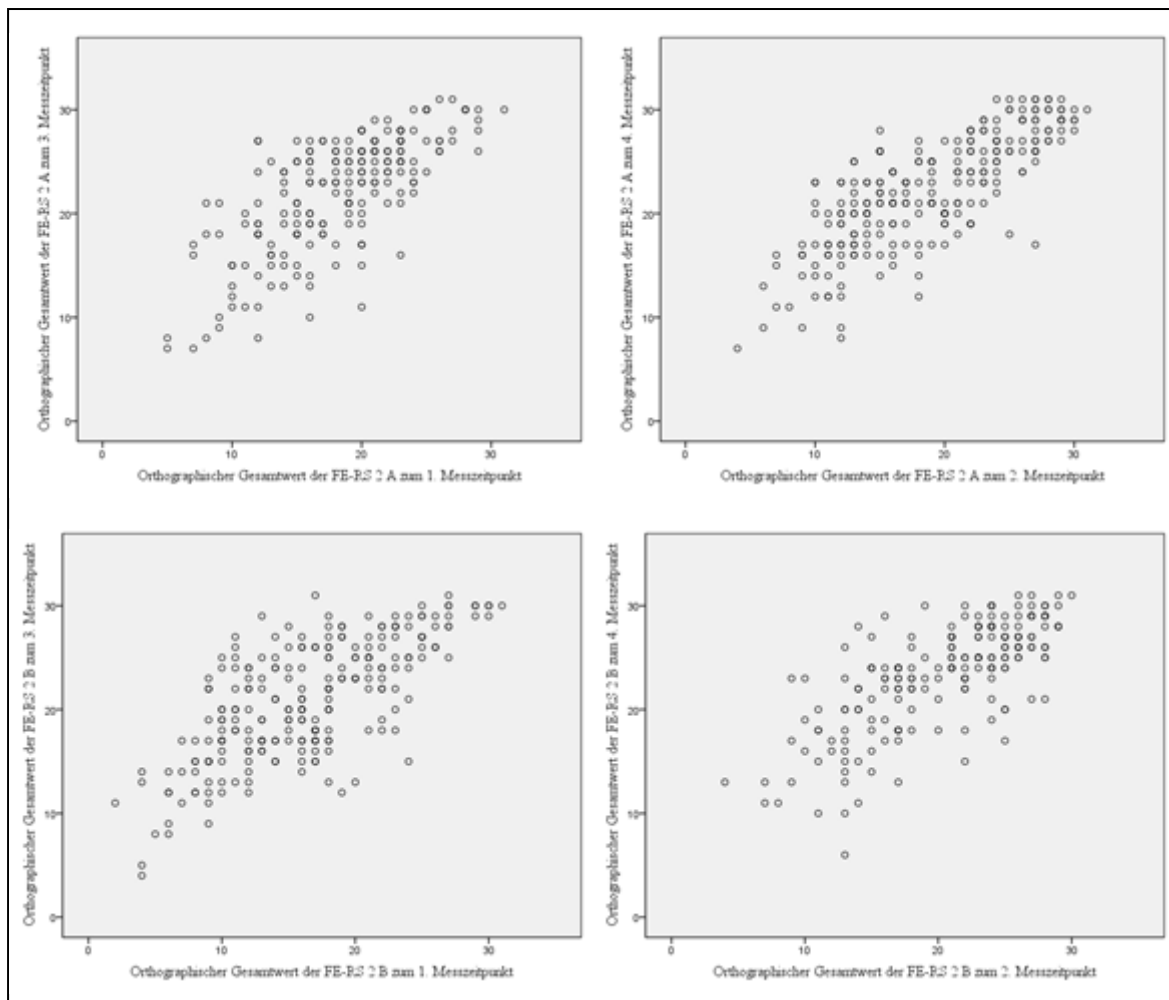


Abbildung X1: Streudiagramme zwischen den Prädiktoren und Kriterien der FE-RS 2
 Anmerkungen: FE-RS 2 A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, FE-RS 2 B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

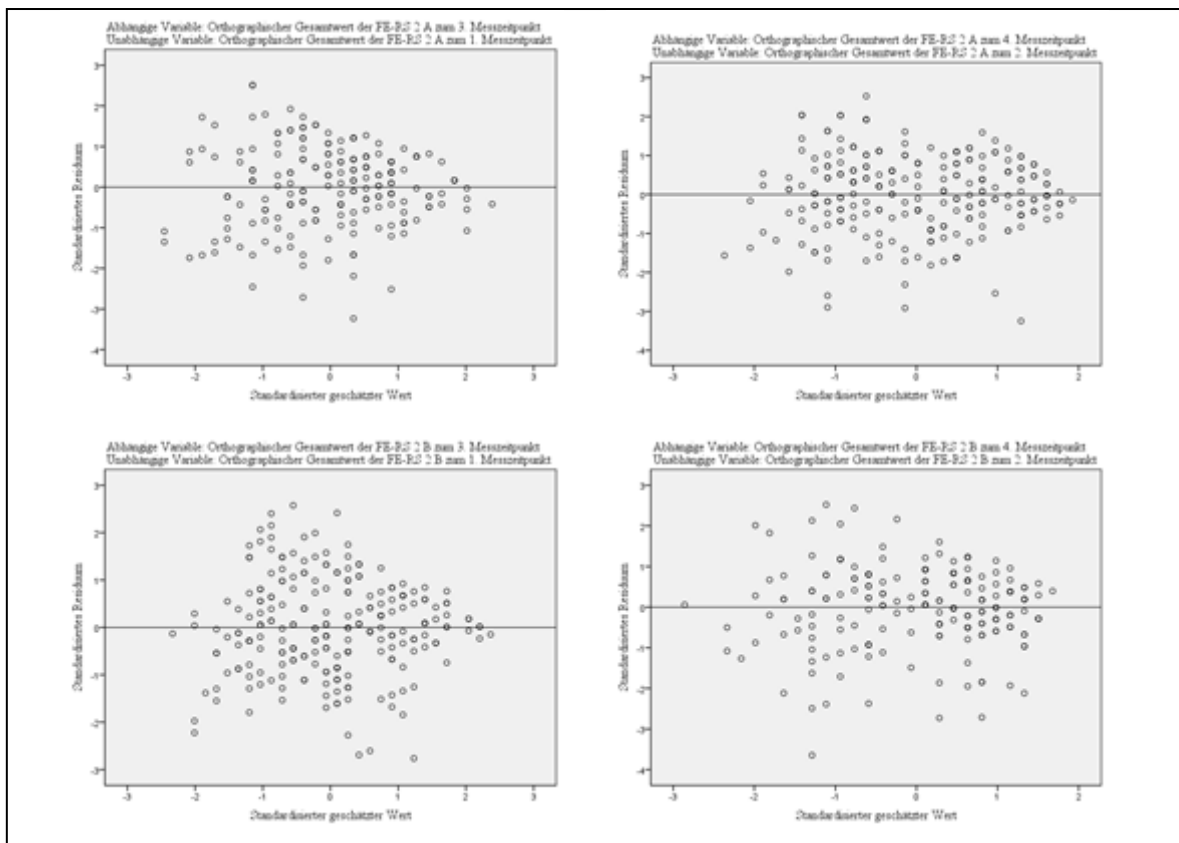


Abbildung X2: Prüfung auf Homoskedastizität der FE-RS 2

Anmerkungen: FE-RS 2 A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, FE-RS 2 B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

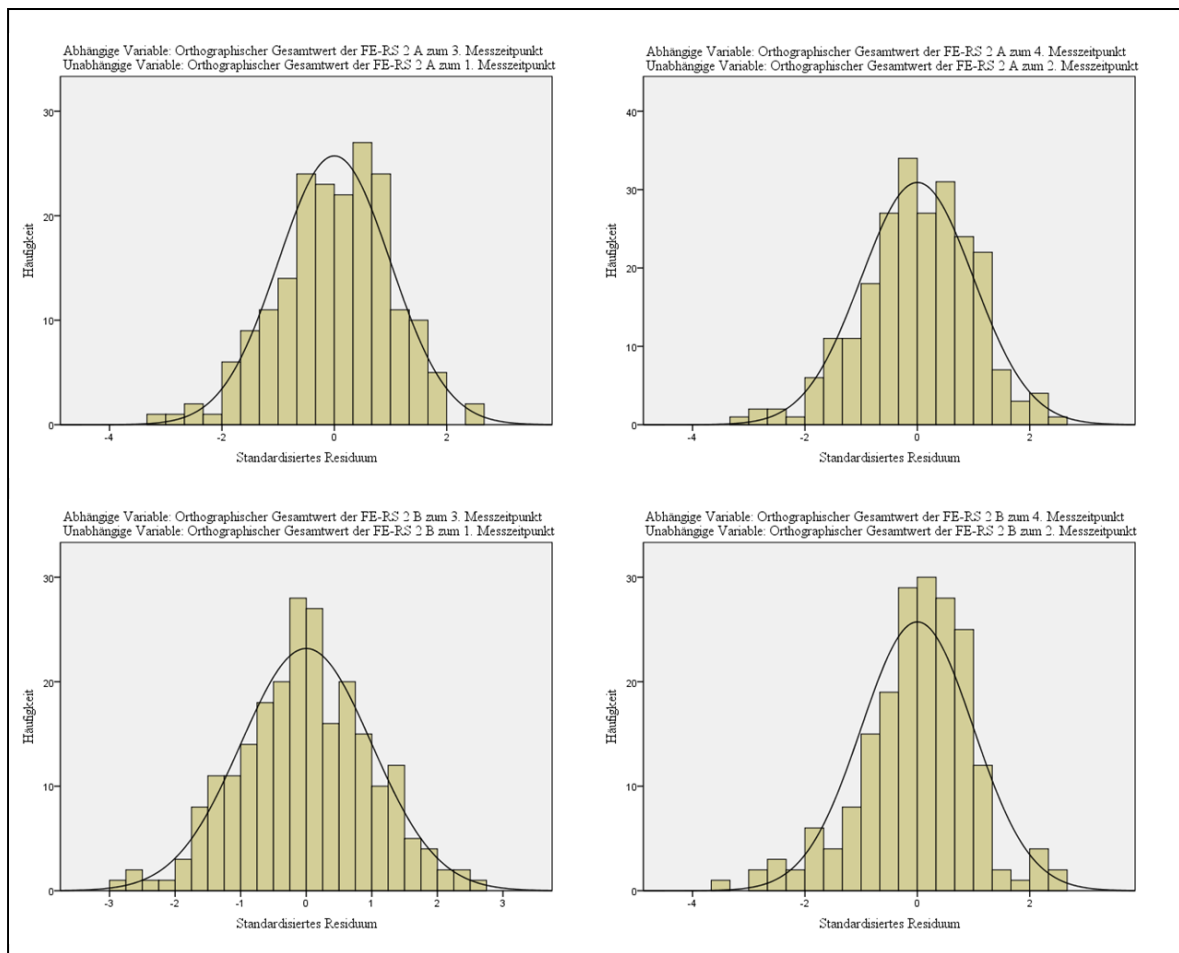


Abbildung X3: Häufigkeitsverteilungen der Residuen der FE-RS 2

Anmerkungen: FE-RS 2 A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, FE-RS 2 B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B

Anhang Y: Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A

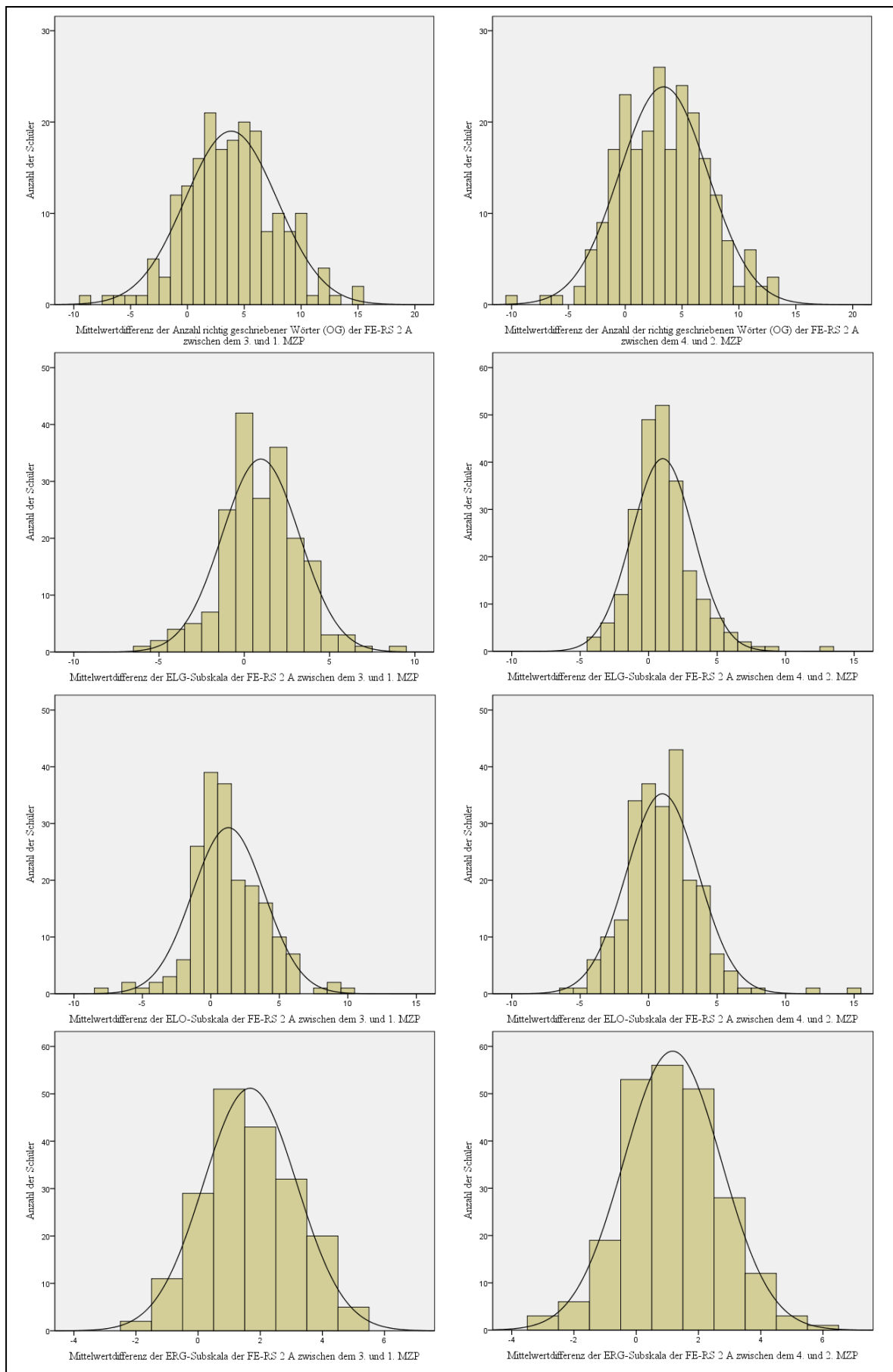


Abbildung Y1: Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form A

Anmerkungen: FE-RS 2 A = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form A, MZP = Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können

Anhang Z: Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der Form B

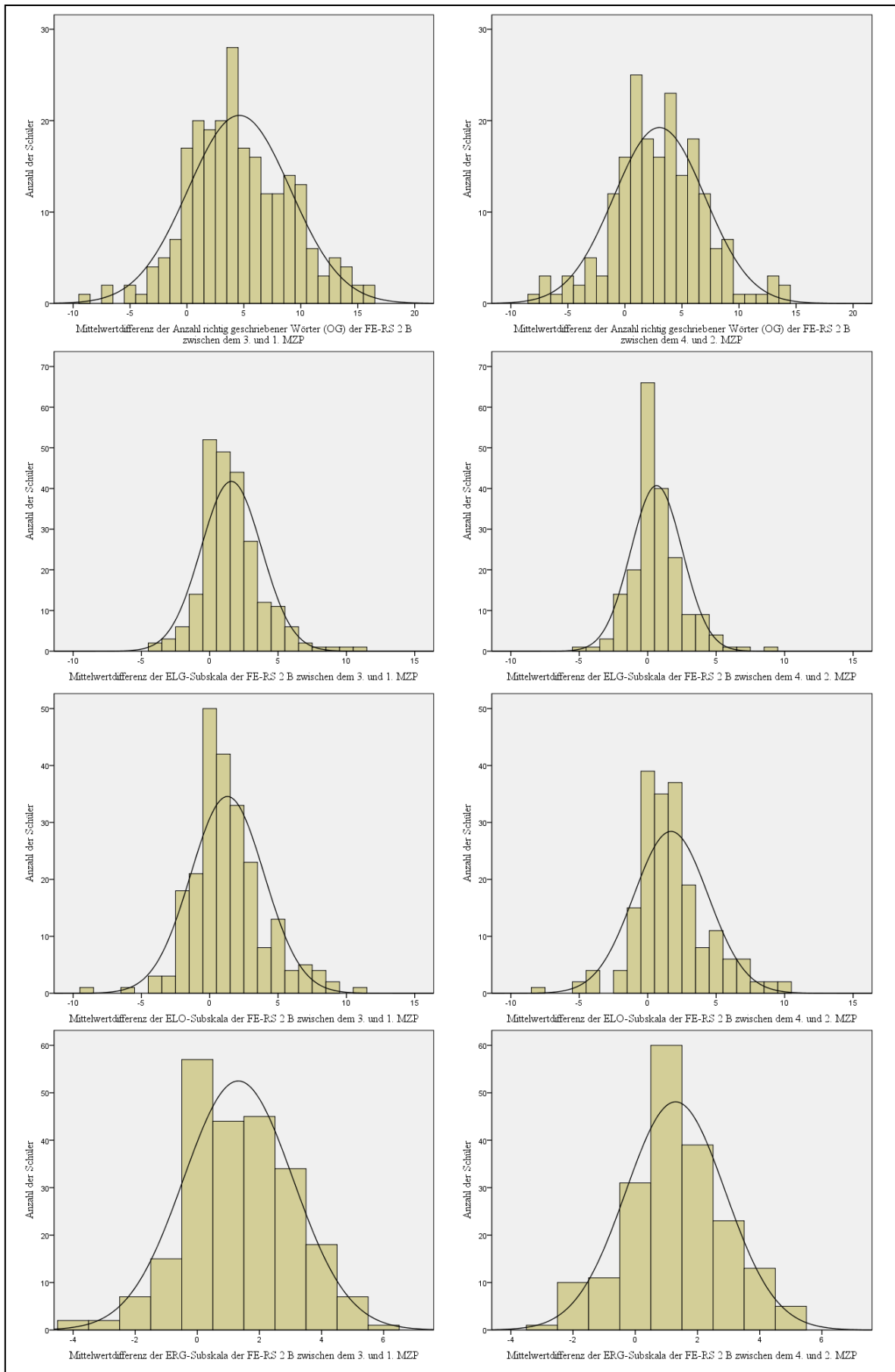


Abbildung Z1: Häufigkeitsverteilungen der Mittelwertdifferenzen des orthographischen Gesamtwerts und der Subskalen der FE-RS 2

Anmerkungen: FE-RS 2 B = Formative Erfassung der Rechtschreibleistung im 2. Schuljahr der Form B, MZP = Messzeitpunkt, OG = orthographischer Gesamtwert, ELG = elementar graphematisches Können, ELO = elementar orthographisches Können, ERG = erweitert graphematisches Können